

Mass spectral molecular networking of living microbial colonies: Supplementary Information

Jeramie Watrous^{ab}, Patrick J. Roach^{c*}, Theodore Alexandrov^{bh}, Brandi S. Heath^c, Jane Y. Yang^{ab}, Roland Kersten,^{ad} Menno vander Voortⁱ, Kit Pogliano^j, Harald Gross^k, Jos M. Raaijmakersⁱ, Bradley S. Moore^{ad}, Julia Laskin^{cl}, Nuno Bandeira^{bfgl} and Pieter C. Dorrestein^{abdl}

^a. Department of Pharmacology and Department of Chemistry and Biochemistry, University of California, San Diego, La Jolla, California, USA.

^b. Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, California, USA.

^c. Chemical and Materials Sciences Division, Pacific Northwest National Laboratory, P.O. Box 999, MSIN K8-88, Richland, WA 99352.

^d. Center for Marine Biotechnology and Biomedicine, Scripps Institution of Oceanography.

^f. Department of Computer science and Engineering, University of California, San Diego, La Jolla, California, USA.

^g. NCRR/UCSD Center for Computational Mass Spectrometry.

^h. Center for Industrial Mathematics, University of Bremen, Bibliothekstr. 1, Bremen, Germany.

ⁱ. Laboratory of Phytopathology, Wageningen University, Droevedaalsesteeg 1, Wageningen PB6708, Netherlands.

^j. Division of biological sciences, University of California, San Diego, La Jolla, California, USA.

^k. Institute for Pharmaceutical Biology, University of Bonn, 53115 Bonn, Germany

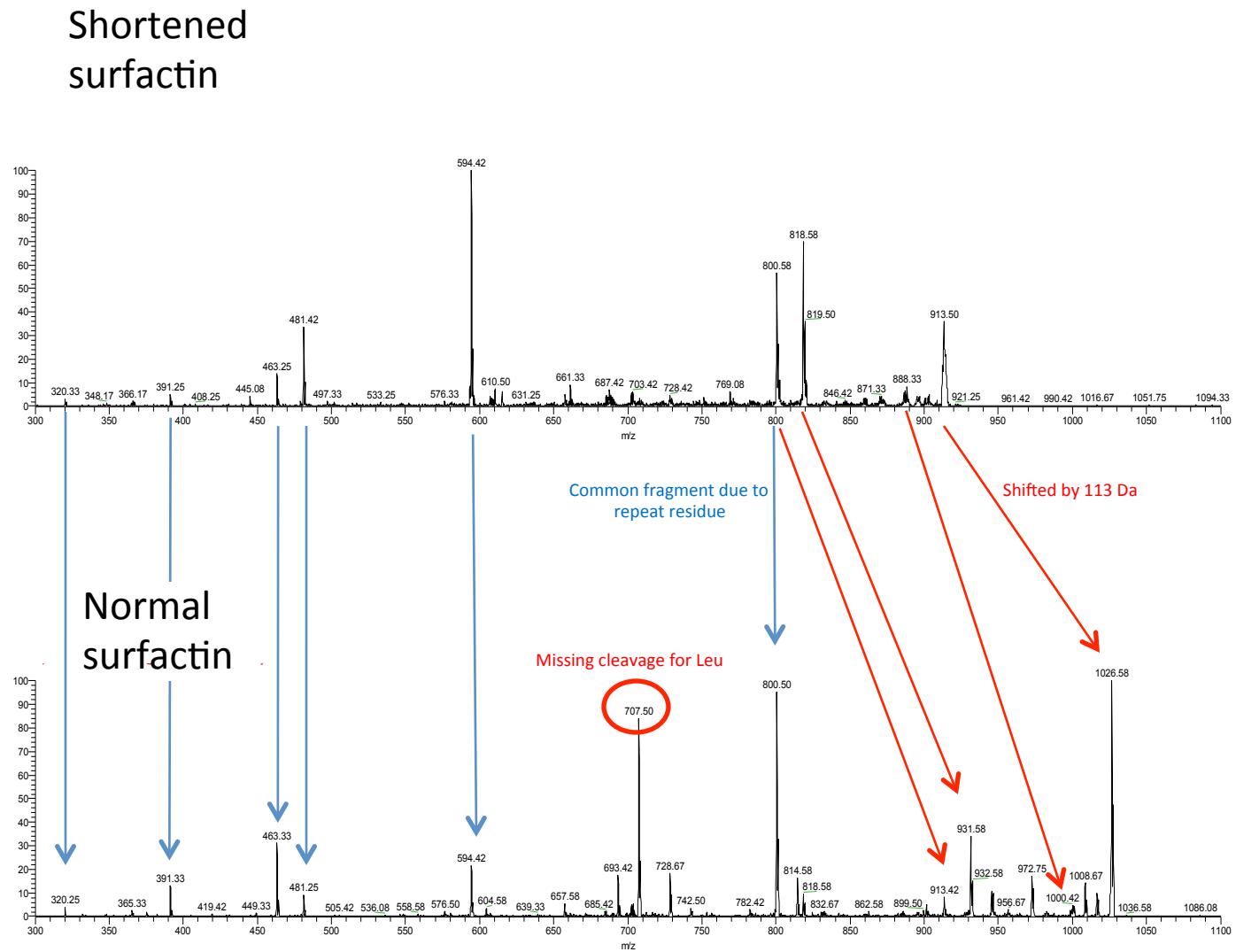
^l. To whom correspondence should be addressed. For the computational aspects of the molecular networks, Nuno Bandeira [bandeira@ucsd.edu], the development of nanoDESI, Julia Laskin [Julia.Laskin@pnnl.gov], mass spectrometric profiling of microbial colonies and interpretation and generation of molecular networks, Pieter C. Dorrestein [pdorrestein@ucsd.edu].

*. Present address, Roach & Associates, 856 North Main Street, Seymour, WI 54165

Table of Contents

Figure S1: Annotation of surfactin less leucine.....	3-5
Figure S2: Proposed biosynthetic skipping mechanism.....	6
Figure S3: Example of nanoDESI solvent bias.....	7
Figure S4: Annotations for observed molecules.....	8-30
Figure S5: Annotation of linear surfactin peptide.....	31-32
Figure S6: Sublancin RNA transcript data.....	33
Figure S7: Plipastatin time course data.....	34
Figure S8: Surfactin time course data.....	35
Figure S9: Prodiginine analogs.....	36-37
Figure S10: Example MS isotopic peak ratios.....	38
Video 1 legend:	39
References	40

Figure S1: Annotated MS/MS spectra for surfactin missing one leucine residue

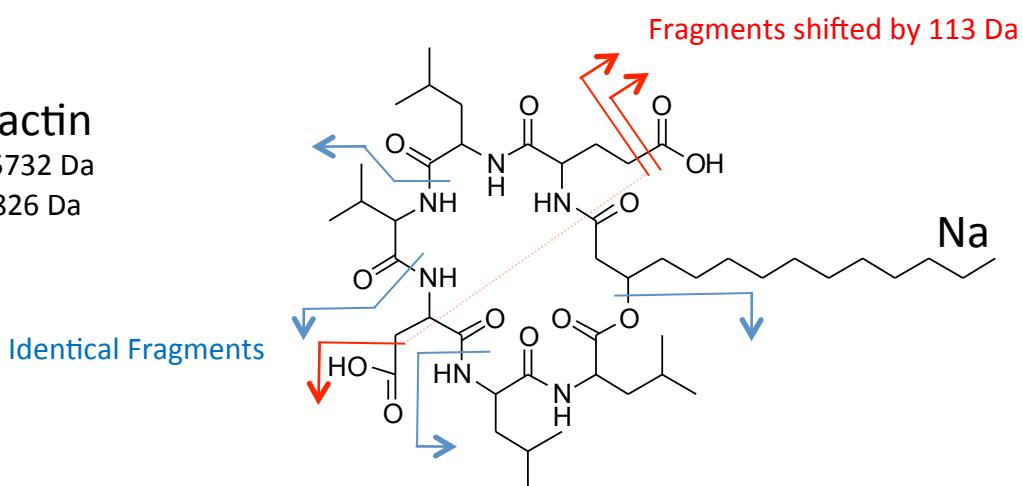


Description: Comparison of the MS^2 CID fragmentation pattern for the surfactin species missing the single leucine residue (top) with the CID fragmentation pattern for the known surfactin compound containing the leucine (bottom) shows identical daughter ions for all fragments on the C-terminal side of the missing leucine (blue arrows) while all daughter ions on the N-terminal side of the missing residue are shifted by the mass of leucine (113 Da...red arrows). Also, the daughter ion for the cleavage between the repeat leucines is missing in the shortened surfactin spectra (red circle).

Figure S1: Annotated MS/MS spectra for surfactin missing one leucine residue

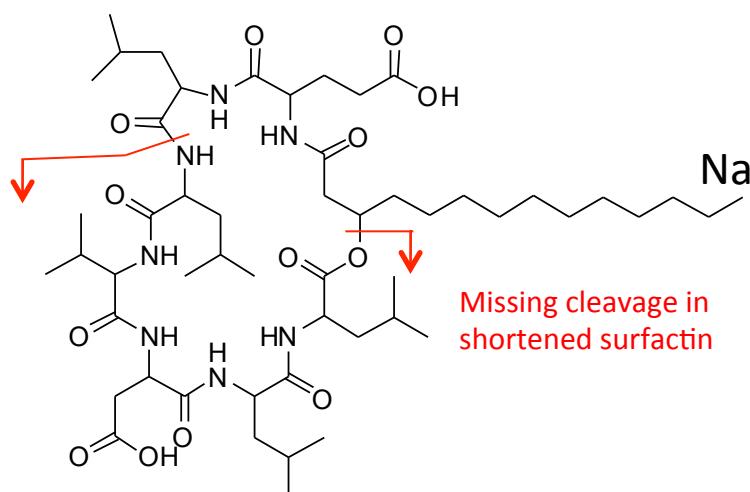
Shortened Surfactin

Calculated Mass: 931.5732 Da
Observed mass: 931.5826 Da



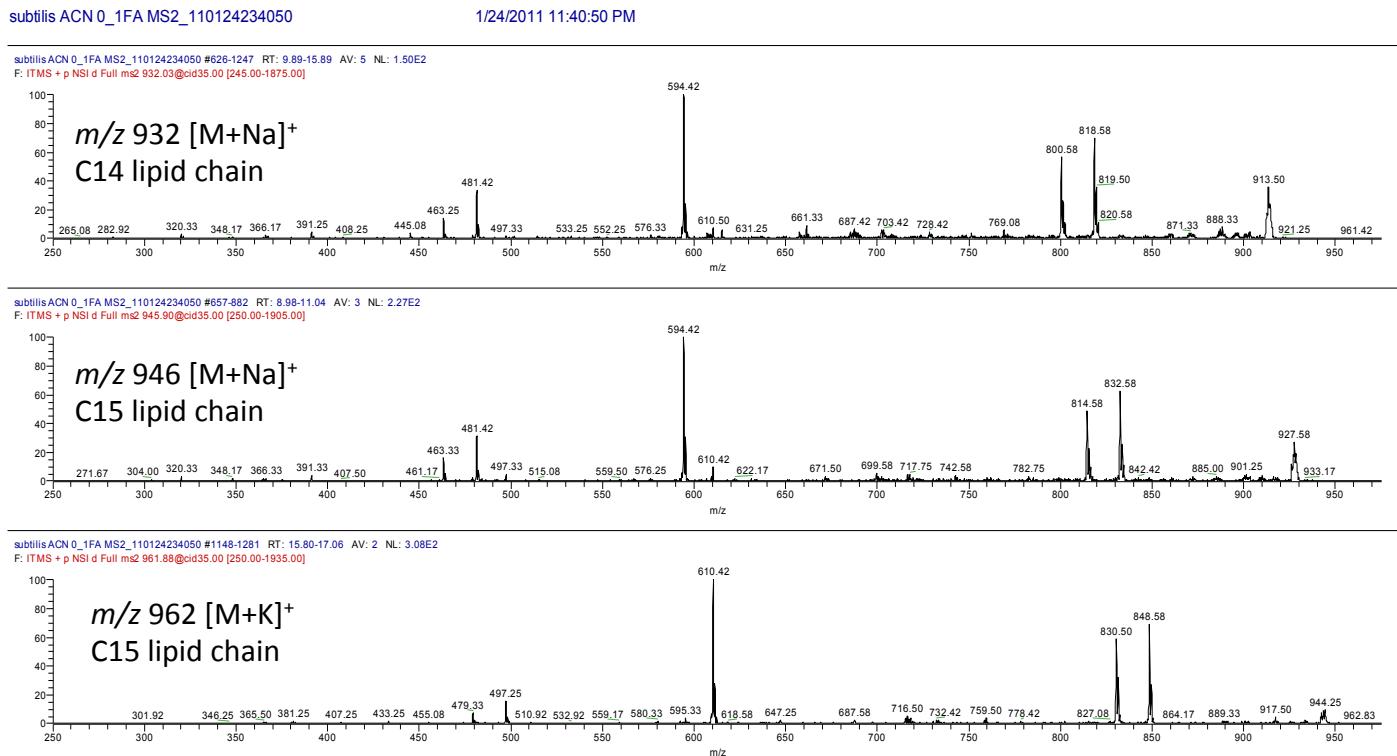
Surfactin

Calculated Mass: 1044.6573 Da



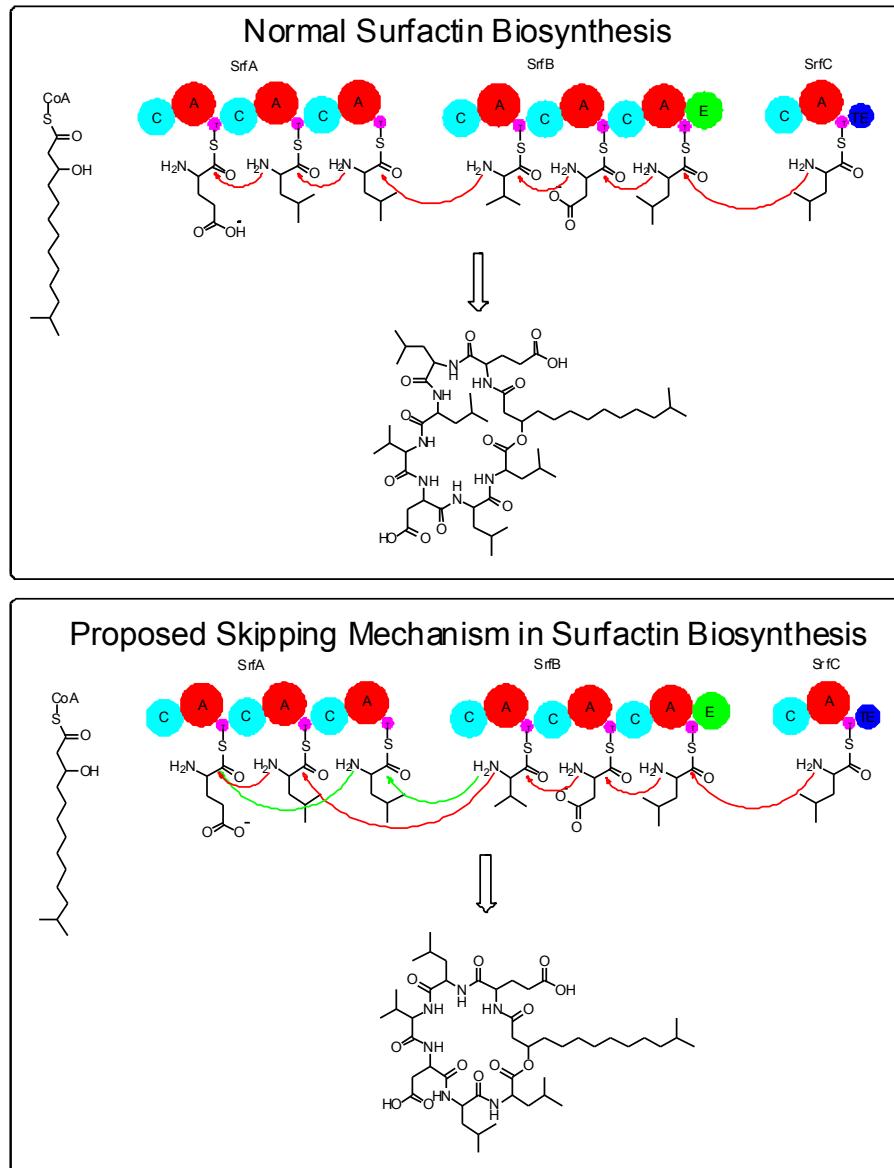
Description: Comparison of the structure for the surfactin species missing the single leucine residue (top) with the structure of the known surfactin compound containing the leucine (bottom) shows identical daughter ions for all fragments on the C-terminal side of the missing leucine (blue arrows) while all daughter ions on the N-terminal side of the missing residue are shifted by the mass of leucine (113 Da...red arrows).

Figure S1: Annotated MS/MS spectra for surfactin missing one leucine residue



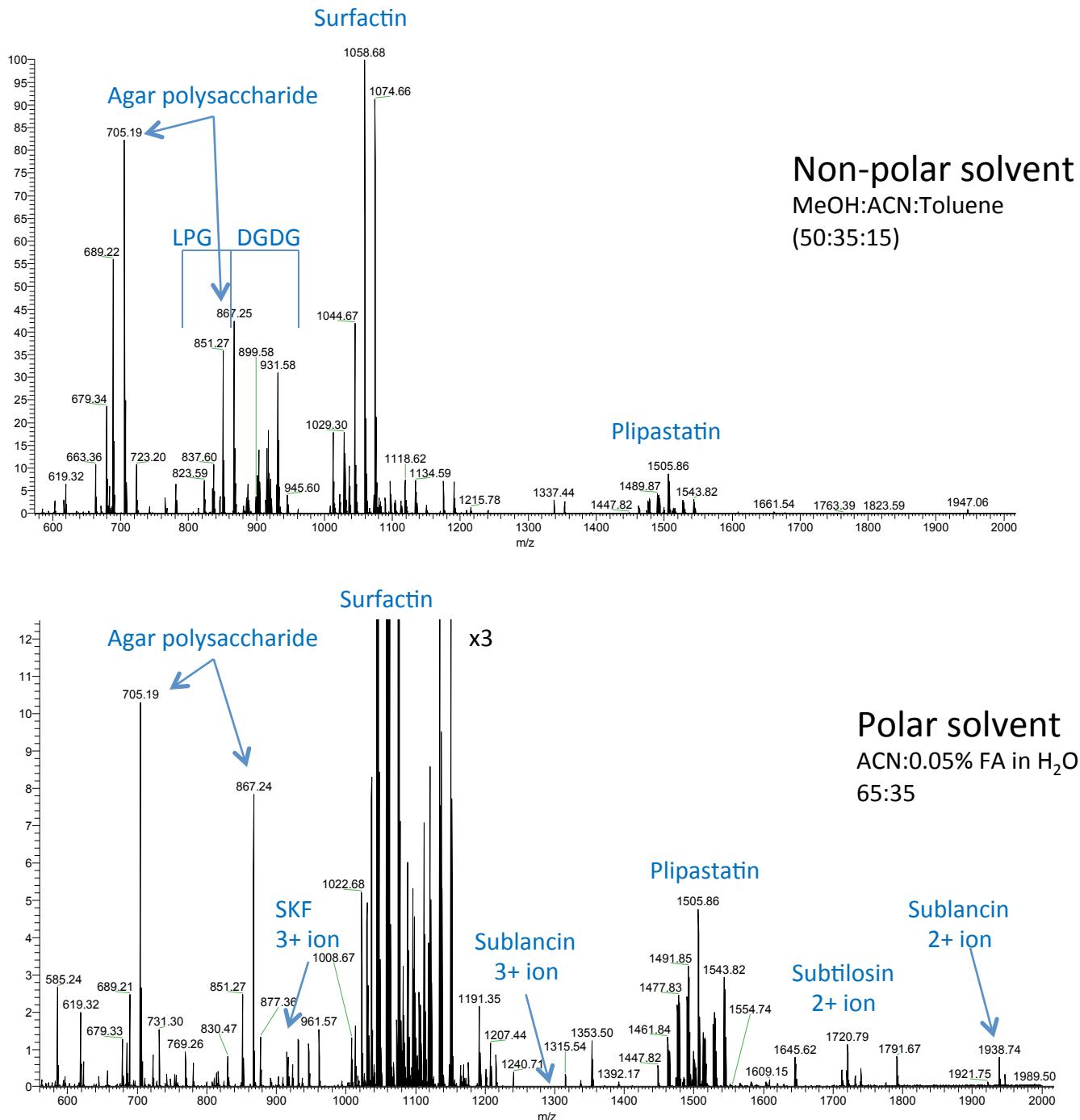
Description: Multiple species of the shortened surfactin compound were observed including C14, C15 and C16 lipid chain lengths with Na and K adducts for each. The above spectra show an example subset of the observed data.

Figure S2: Proposed biosynthetic mechanism for skipping of a leucine residue during surfactin production



Description: The top mechanism shows the typical surfactin biosynthesis while the bottom figure is a proposed mechanism to possibly explain how the biosynthetic machinery can skip one of the repeat leucine residues during production. Here we hypothesize that it may be possible for SrfB to receive the growing lipopeptide from SrfA after either leucine has been added in module 1 or 2 or that the product from module 1 can skip module 2 and go straight to module 3 (green arrows).

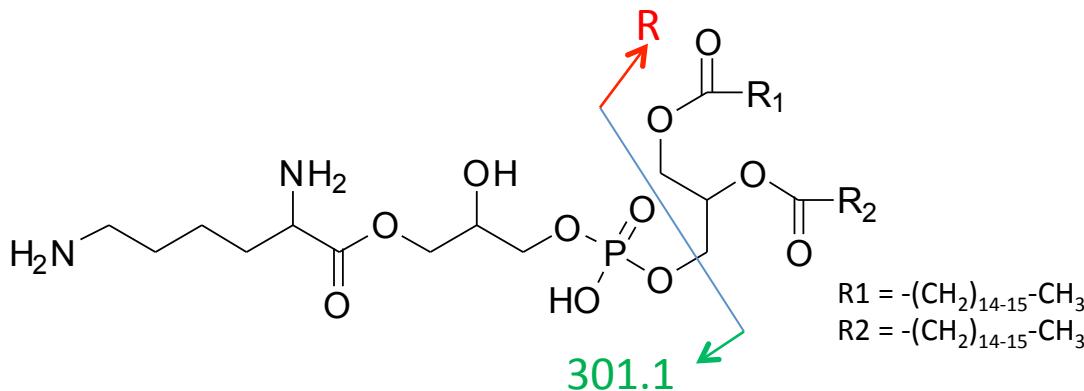
Figure S3: nanoDESI MS1 spectra from *Bacillus subtilis* 3610 using polar and non-polar solvent



Description: Example MS1 Orbitrap positive mode spectra for *B. subtilis* 3610 using a non-polar (top) and polar (bottom) solvent. While many compounds are soluble in both solvents, certain compounds are unique to each condition such as the lipid families LPG and DGDG in the non-polar solvent and the macrocyclic peptides sublancin, SKF and subtilosin in the polar solvent.

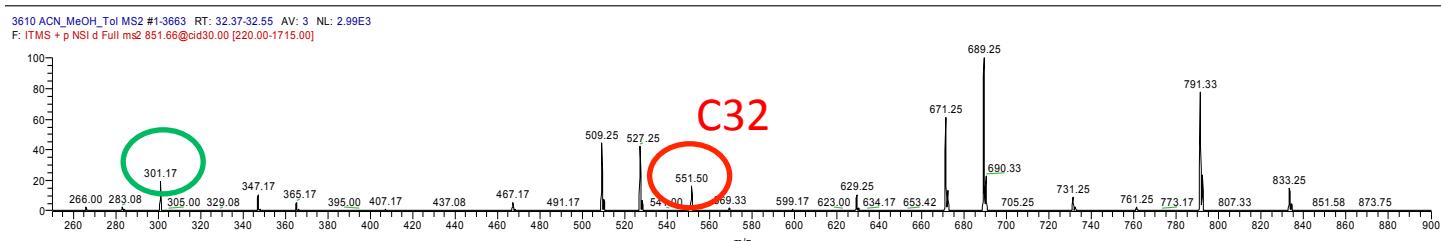
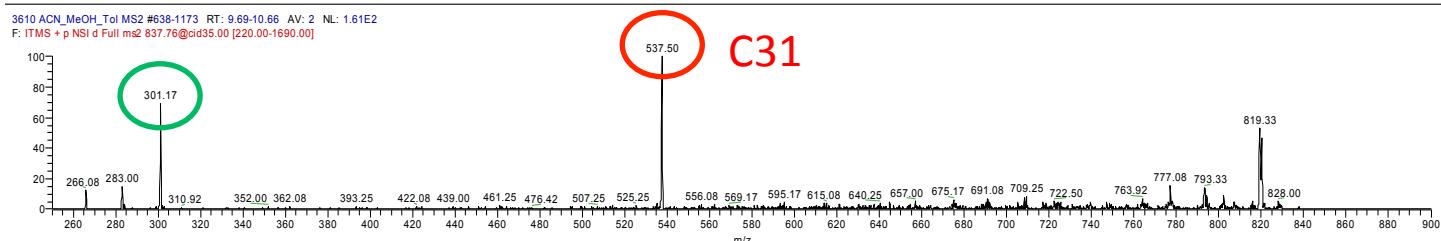
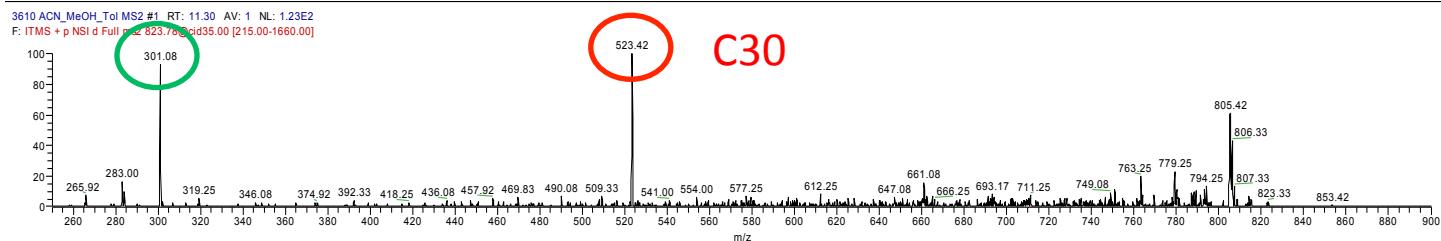
Figure S4: Annotation of MS/MS spectra for reported compounds

Bacillus subtilis 3610: Lysyl-phosphatidylglycerol (LPG)



3610 ACN_MeOH_Tol MS2

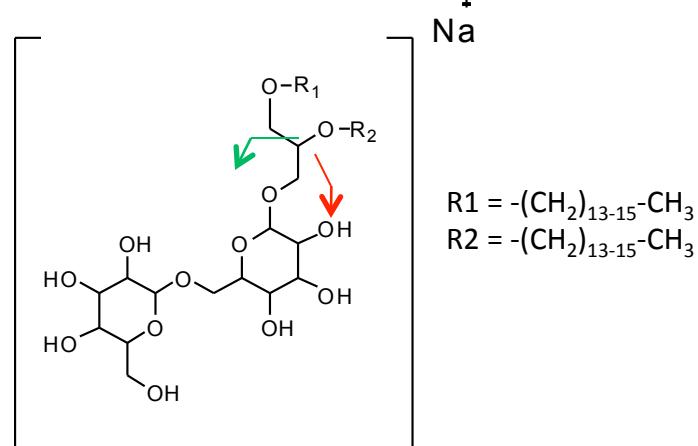
1/24/2011 8:36:43 PM



Description: Example subset of MS2 CID spectra for the LPG family of compounds. All LPG compounds exhibited a common loss of the head group at m/z 301 along with a second major fragment containing the lipid chains. The bottom spectrum contains extra peaks due to one of the agar polysaccharide peaks being within the mass isolation window. The spectra are consistent with published MS2 data.²

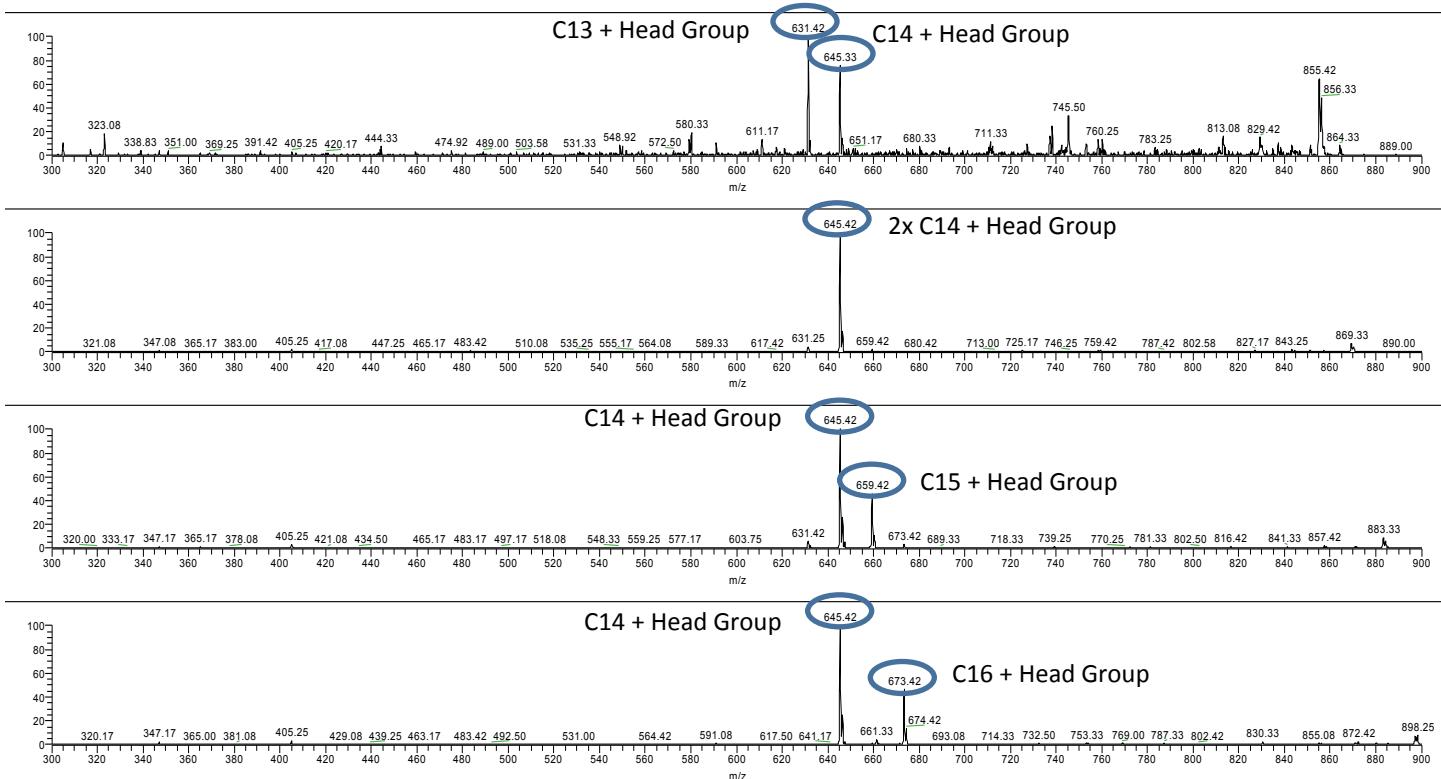
Figure S4: Annotation of MS/MS spectra for reported compounds

Bacillus subtilis 3610: Diglycosyl diglycerides (DGDG)



3610 ACN_MeOH_Tol MS2

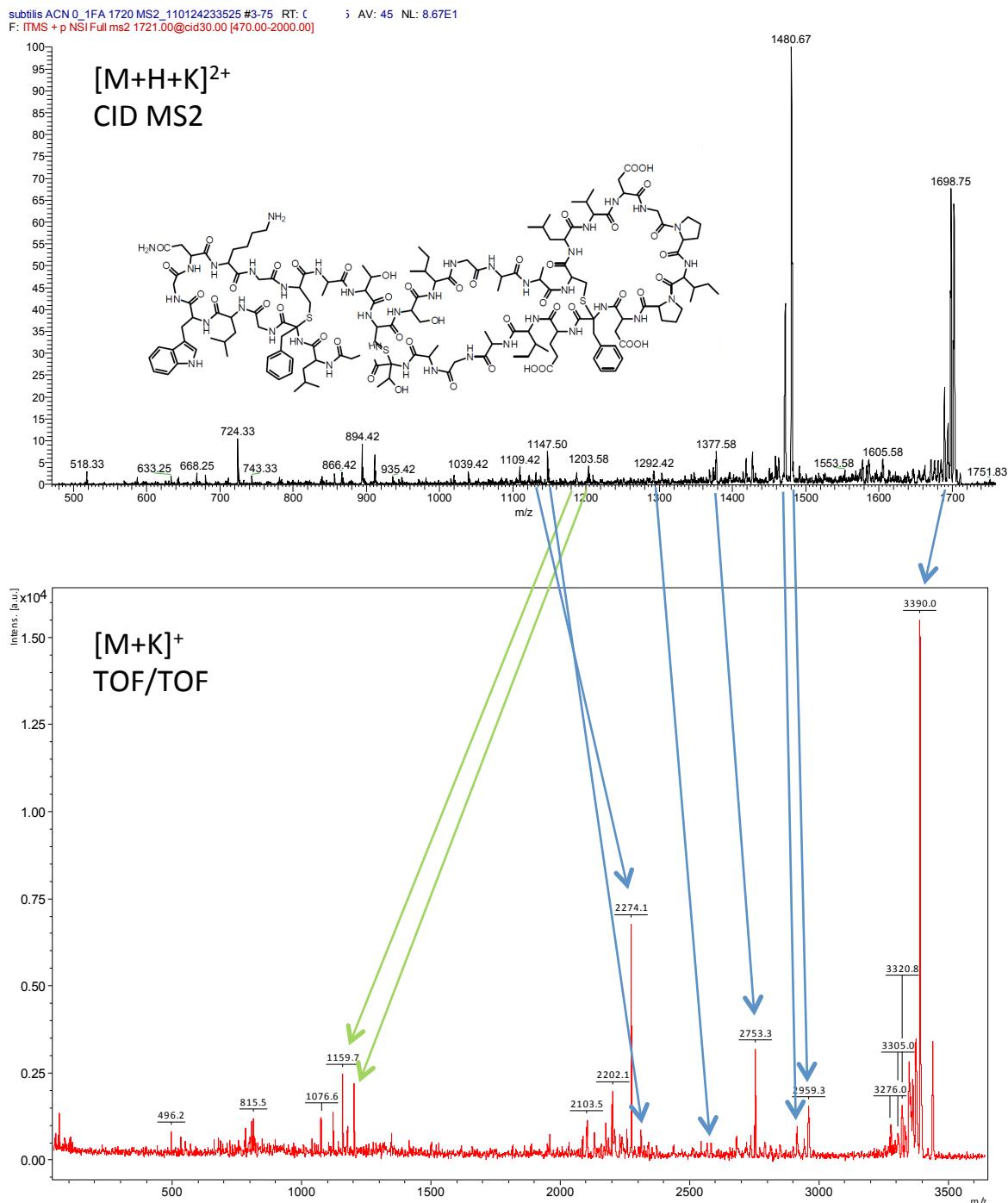
1/24/2011 8:36:43 PM



Description: Example subset of MS2 CID spectra for the DGDG family of compounds. Unlike the LPG compounds, the head group is retained in the mass fragments where each of the 2 fragments show the head group plus one of the two lipid chains. The spectra are consistent with published MS2 data.²

Figure S4: Annotation of MS/MS spectra for reported compounds

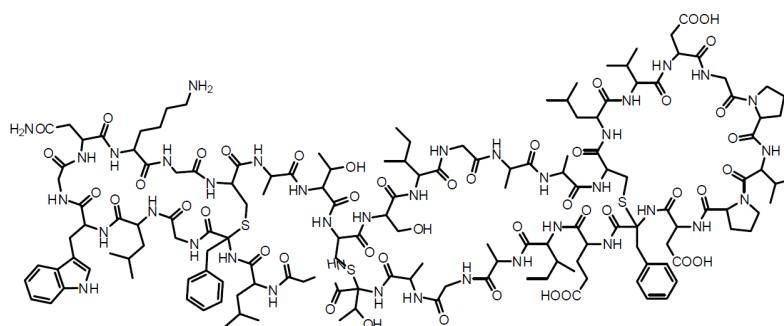
Bacillus subtilis 3610: Subtilisin



Description: Confirmation of subtilisin from nanoDESI experiments (top) by comparing with TOF/TOF MALDI fragmentation pattern for purified subtilisin (bottom). Blue arrows indicate a 2+ ion in the nanoDESI CID spectra corresponding to a 1+ ion in the TOF/TOF while green arrows indicate both are 1+.

Figure S4: Annotation of MS/MS spectra for reported compounds

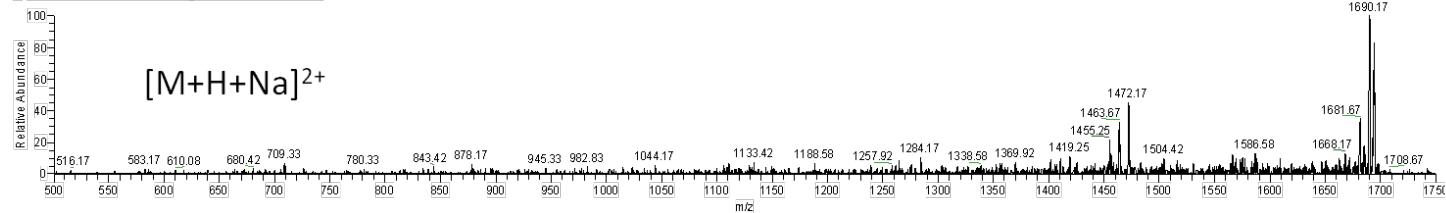
Bacillus subtilis 3610: Subtilisin



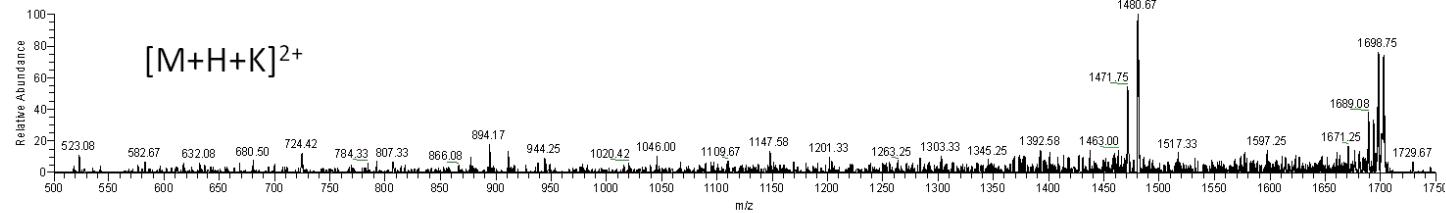
subtilis ACN 0_1FA MS2_110124234050

1/24/2011 11:40:50 PM

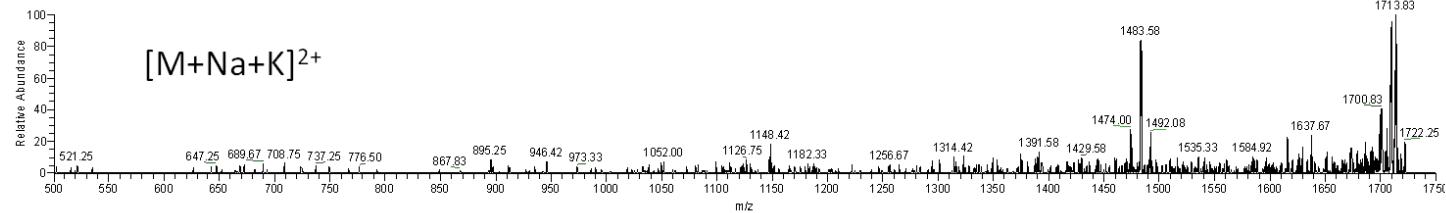
subtilis ACN 0_1FA MS2_110124234050 #980-1105 RT: 13.38-13.75 AV: 2 NL: 6.83E1
F: ITMS + p NSI d Full ms2 1712.63@dd30.00 [460.00-2000.00]



subtilis ACN 0_1FA MS2_110124234050 #436-673 RT: 6.63-7.61 AV: 2 NL: 5.18E1
F: ITMS + p NSI d Full ms2 1720.74@dd35.00 [460.00-2000.00]



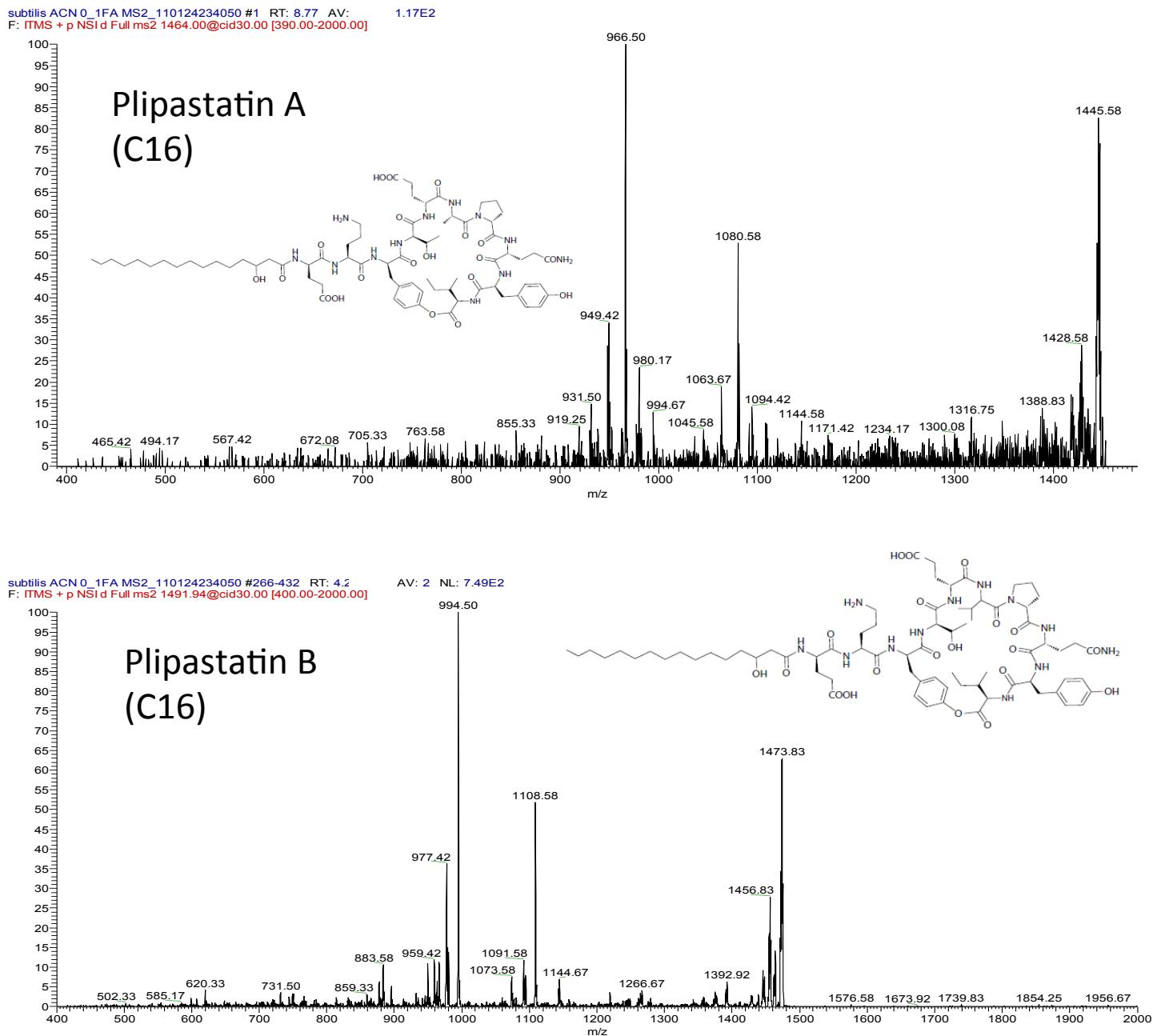
subtilis ACN 0_1FA MS2_110124234050 #1 RT: 14.12 AV: 1 NL: 7.00E1
F: ITMS + p NSI d Full ms2 1732.00@dd30.00 [465.00-2000.00]



Description: Example subset of MS2 CID spectra for the subtilisin peak cluster. Poor fragmentation is characteristic of cyclic peptides containing disulfide bonds.

Figure S4: Annotation of MS/MS spectra for reported compounds

Bacillus subtilis 3610: Plipastatin



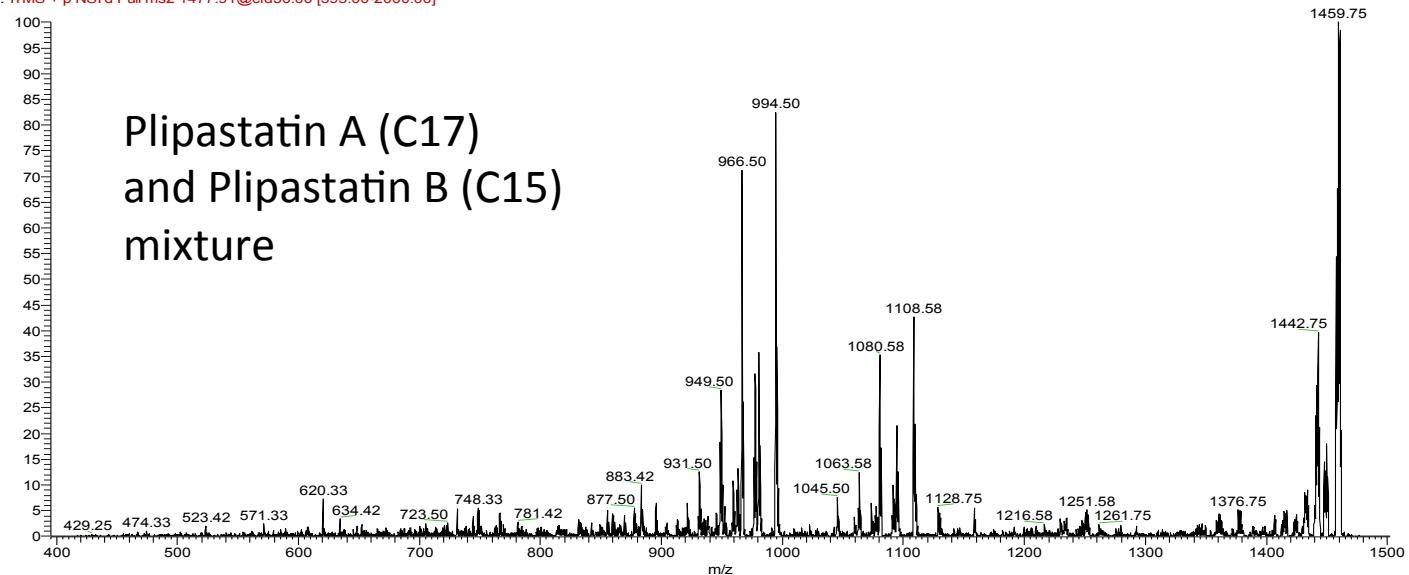
Description: Example subset of MS2 CID spectra for the plipastatin family of compounds. The spectra are consistent with published MS2 spectra.³

Figure S4: Annotation of MS/MS spectra for reported compounds

Bacillus subtilis 3610: Plipastatin

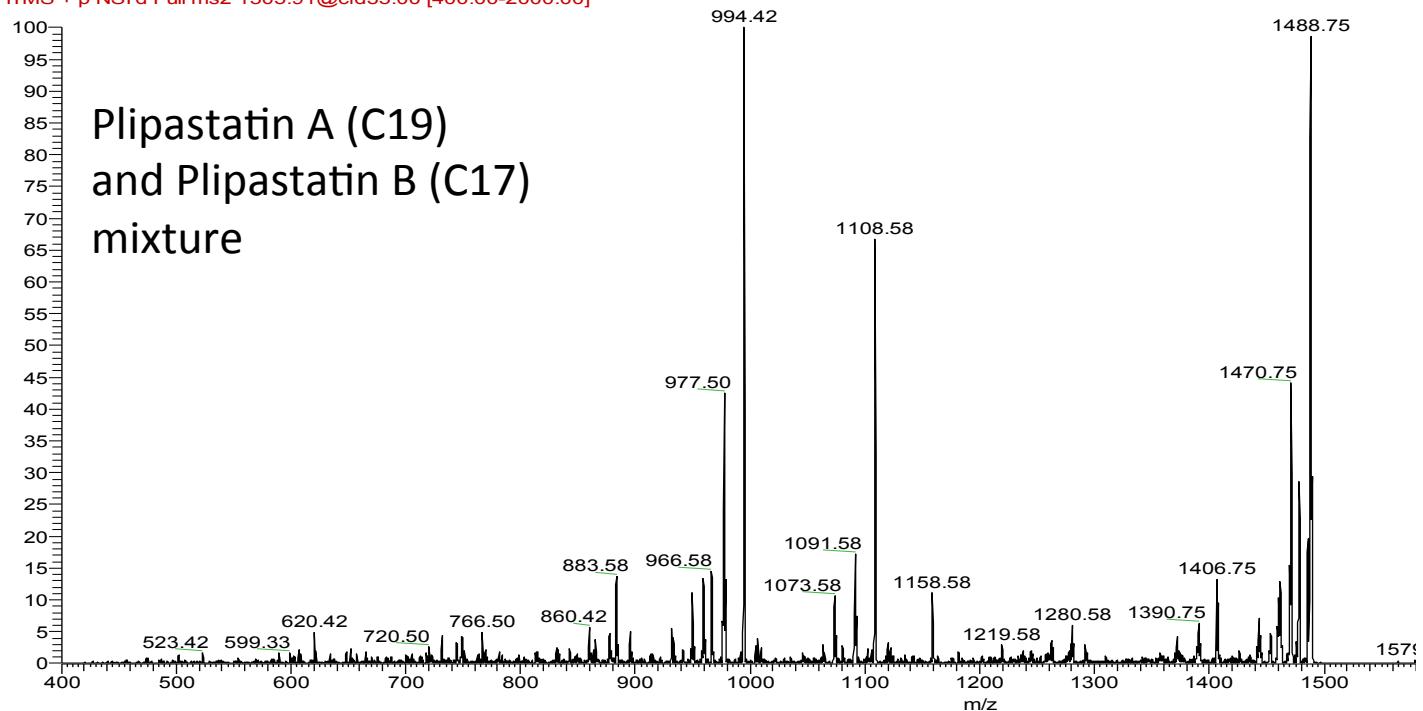
subtilis ACN 0_1FA MS2_110124234050 #442-554 RT: 6.3
F: ITMS + p NSI d Full ms2 1477.91@cid30.00 [395.00-2000.00]

AV: 3 NL: 4.92E2



subtilis ACN 0_1FA MS2_110124234050 #117-580 RT: 2.3
F: ITMS + p NSI d Full ms2 1505.91@cid35.00 [400.00-2000.00]

AV: 5 NL: 3.98E2

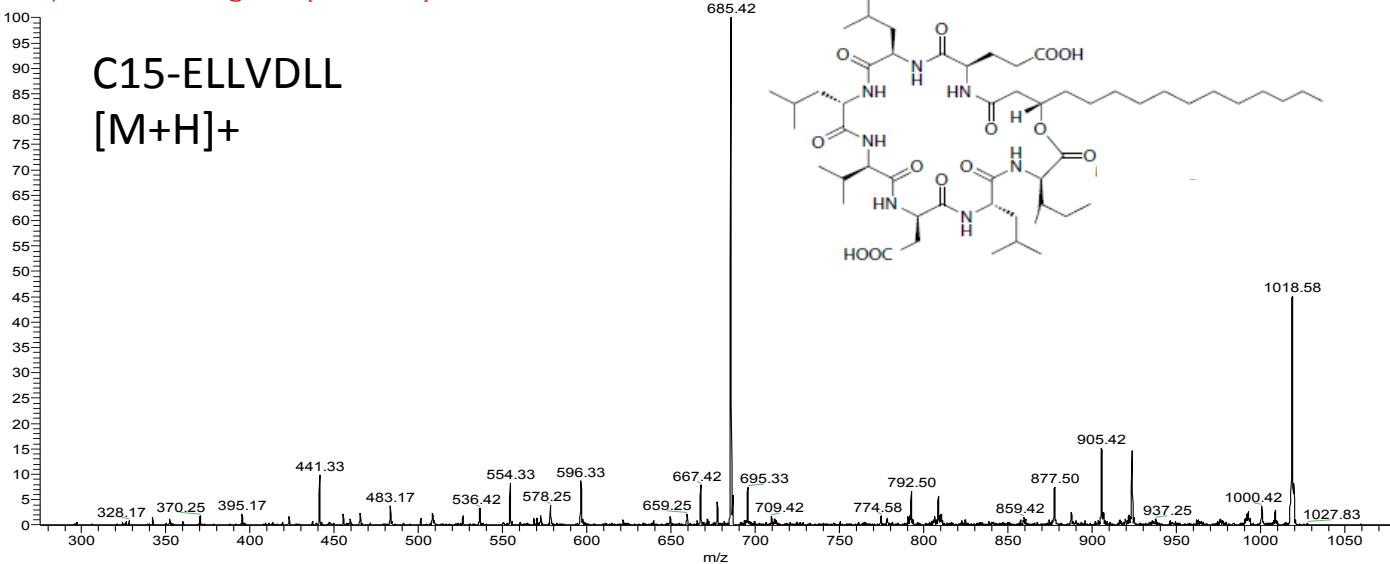


Description: Example subset of MS2 CID spectra for the plipastatin family of compounds. The spectra are consistent with published MS2 spectra.³

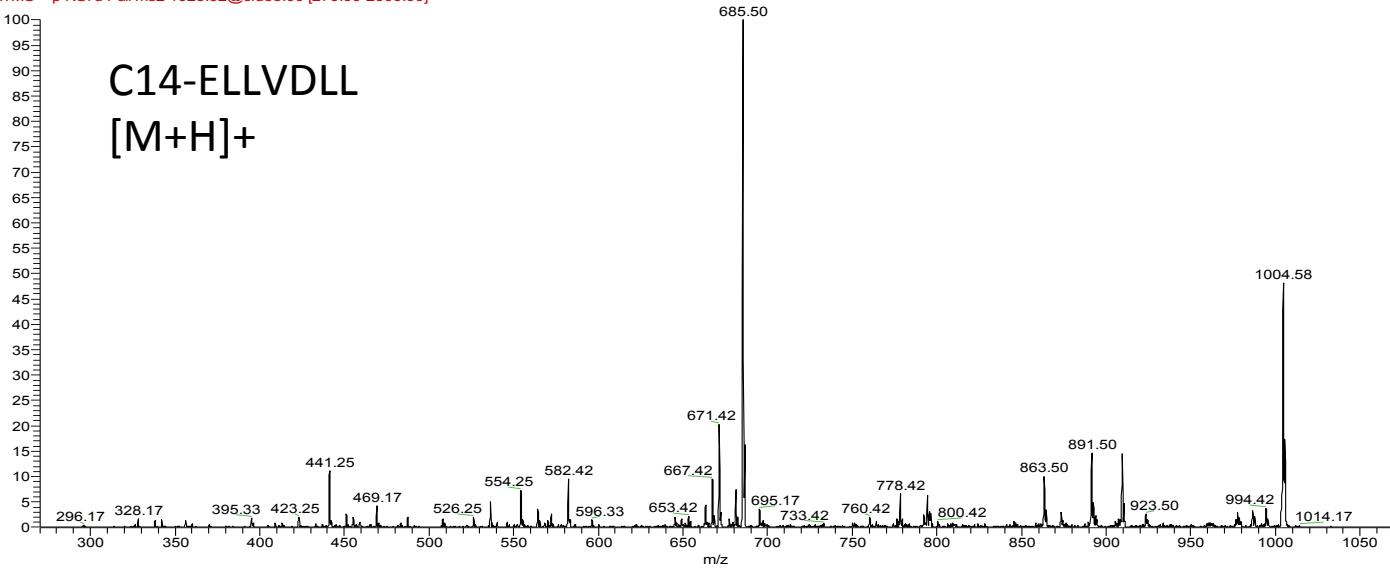
Figure S4: Annotation of MS/MS spectra for reported compounds

Bacillus subtilis 3610: Surfactin

subtilis ACN 0_1FA MS2_110124234050 #73-391 RT: 1.7^E
IV: 3 NL: 7.37E2
F: ITMS + p NSId Full ms2 1036.98@cid35.00 [275.00-2000.00]

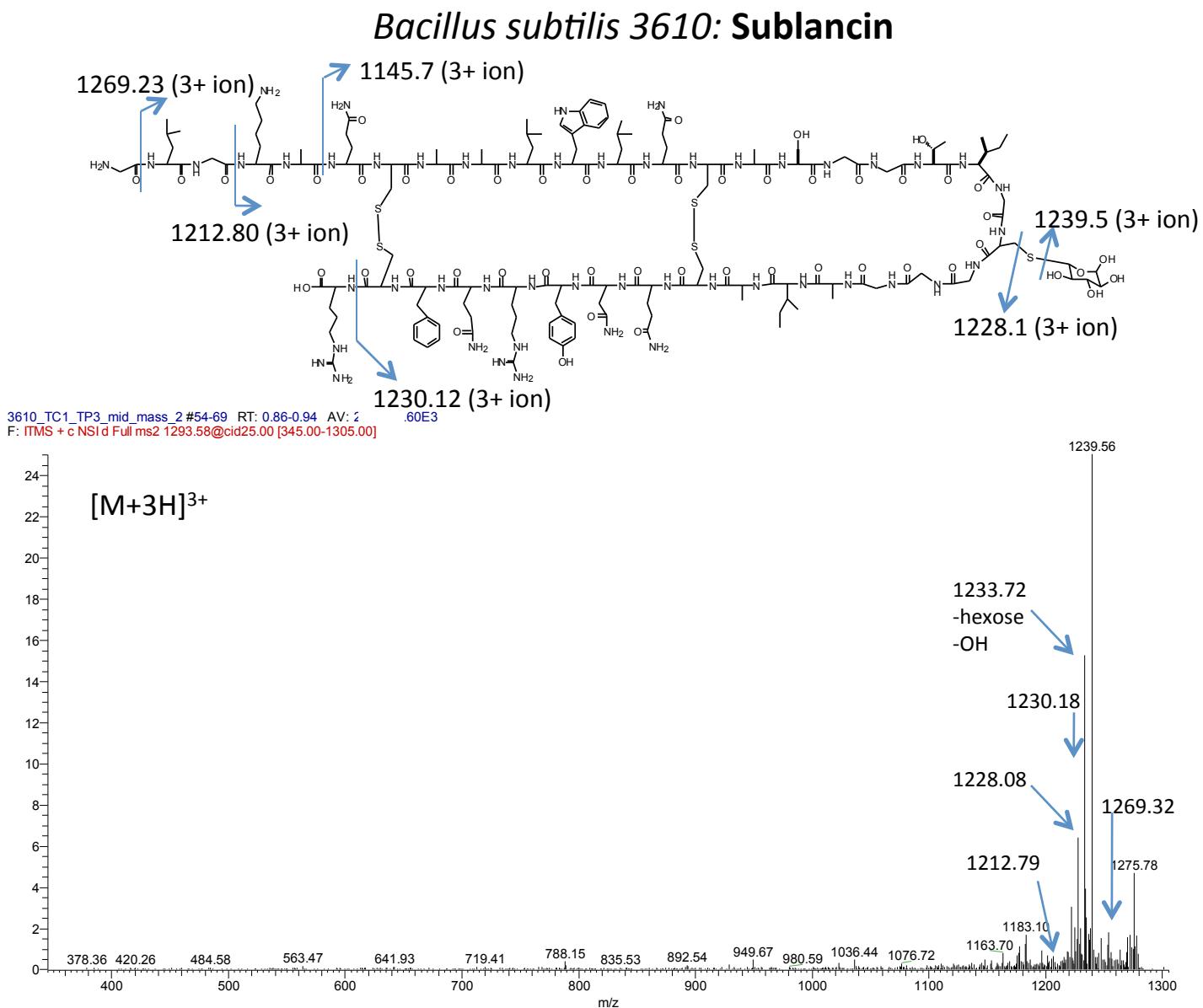


subtilis ACN 0_1FA MS2_110124234050 #84-321 RT: 2.1^E
IV: 4 NL: 5.33E2
F: ITMS + p NSId Full ms2 1023.02@cid35.00 [270.00-2000.00]



Description: Example subset of MS2 CID spectra for the surfactin family of compounds. The spectra are consistent with published MS2 spectra.¹

Figure S4: Annotation of MS/MS spectra for reported compounds

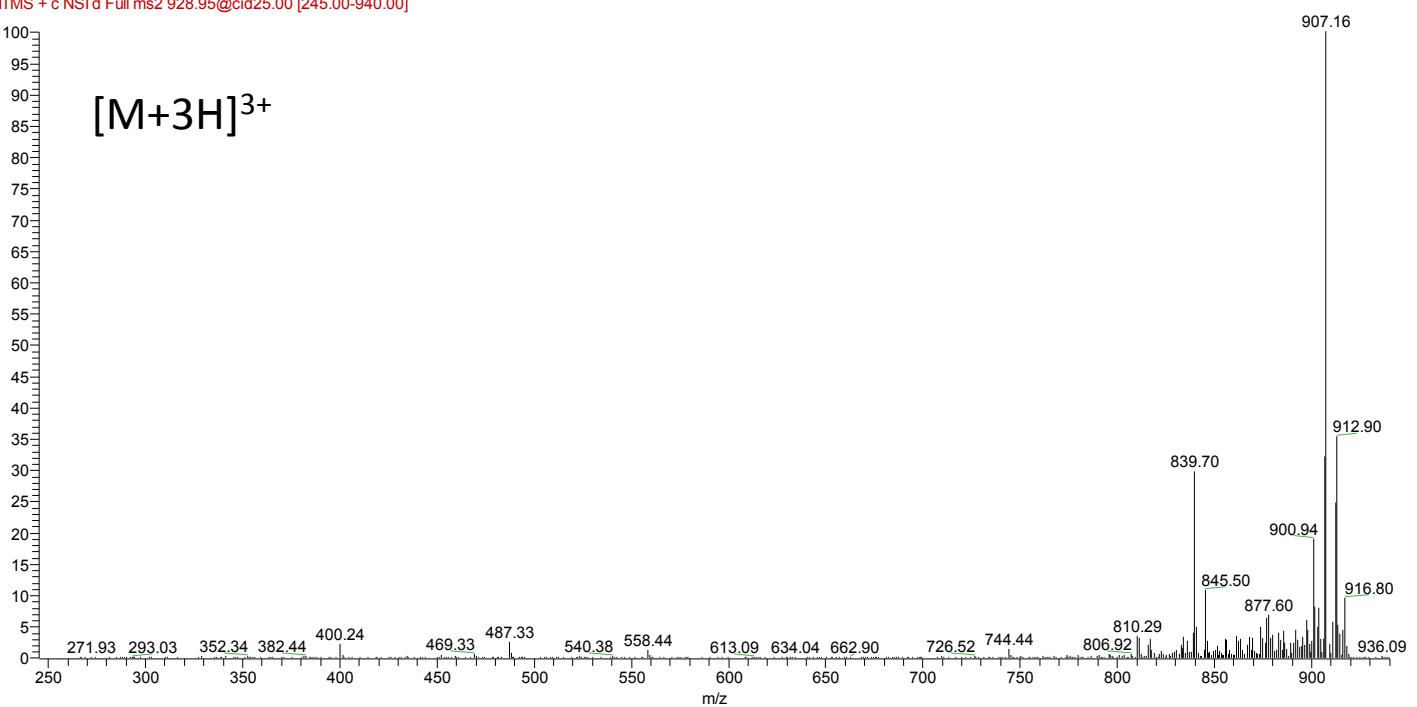
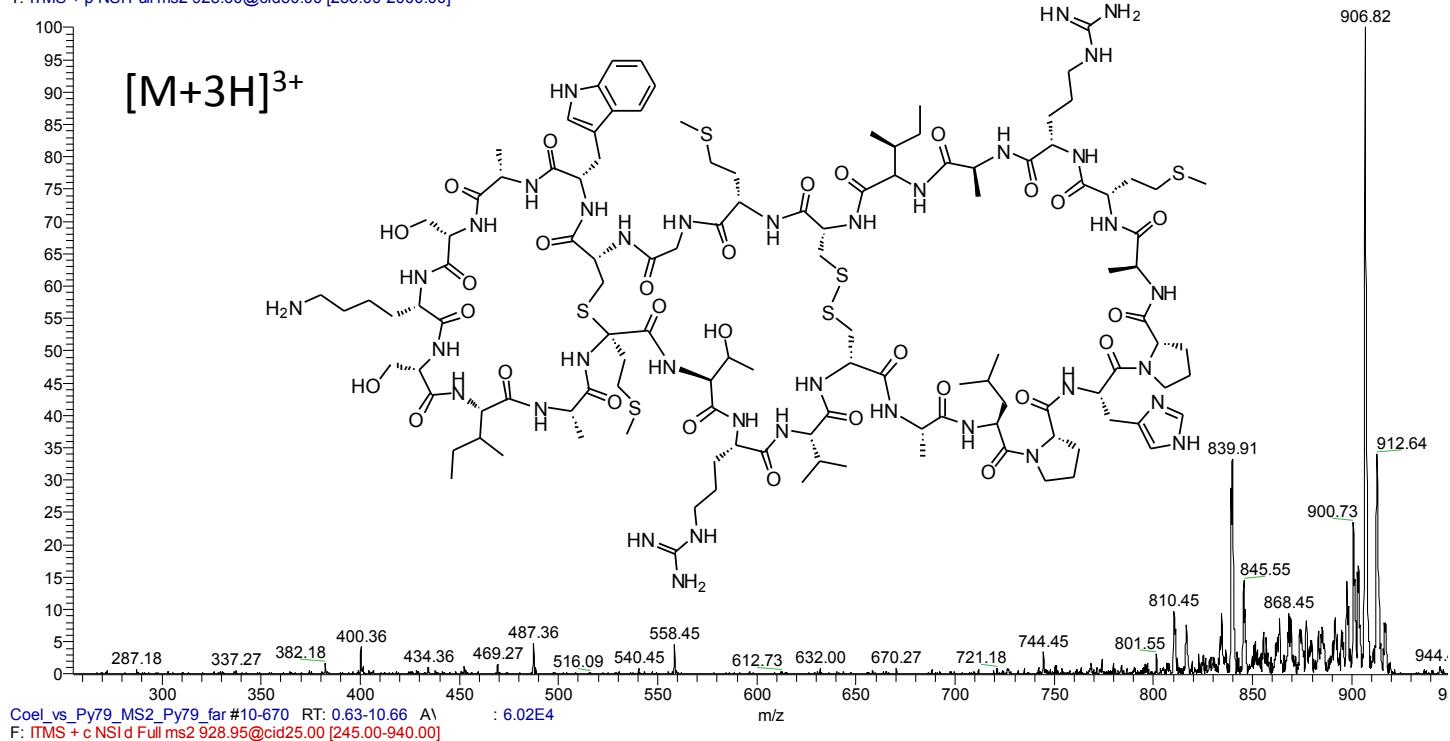


Description: CID MS2 spectrum for sublancin $[M+3H]^{3+}$ ion. The only major fragment observed is the loss of the sugar (m/z 1239.56), which is consistent with the published MS2 spectra for the native glycopeptide.⁴

Figure S4: Annotation of MS/MS spectra for reported compounds

Bacillus subtilis PY79: SKF

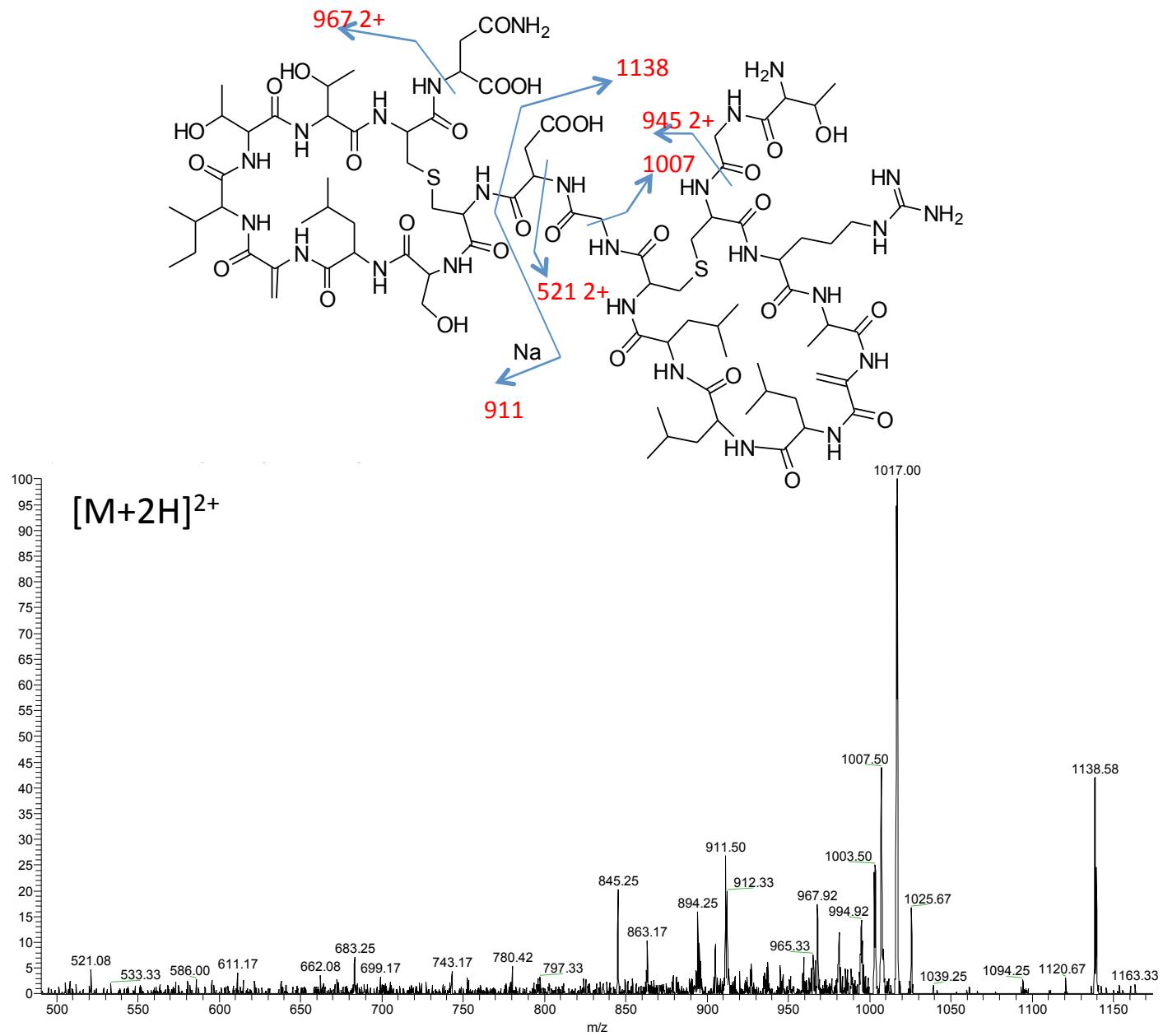
2782ITMS2 #1 RT: 0.00 AV: 1 NL: 4.60E5
T: ITMS + p NSI Full ms2 928.60@cid30.00 [255.00-2000.00]



Description: CID MS2 of the [M+3H]³⁺ ion of SKF (bottom). Spectra is consistent with purified SKF (Top).⁵

Figure S4: Annotation of MS/MS spectra for reported compounds

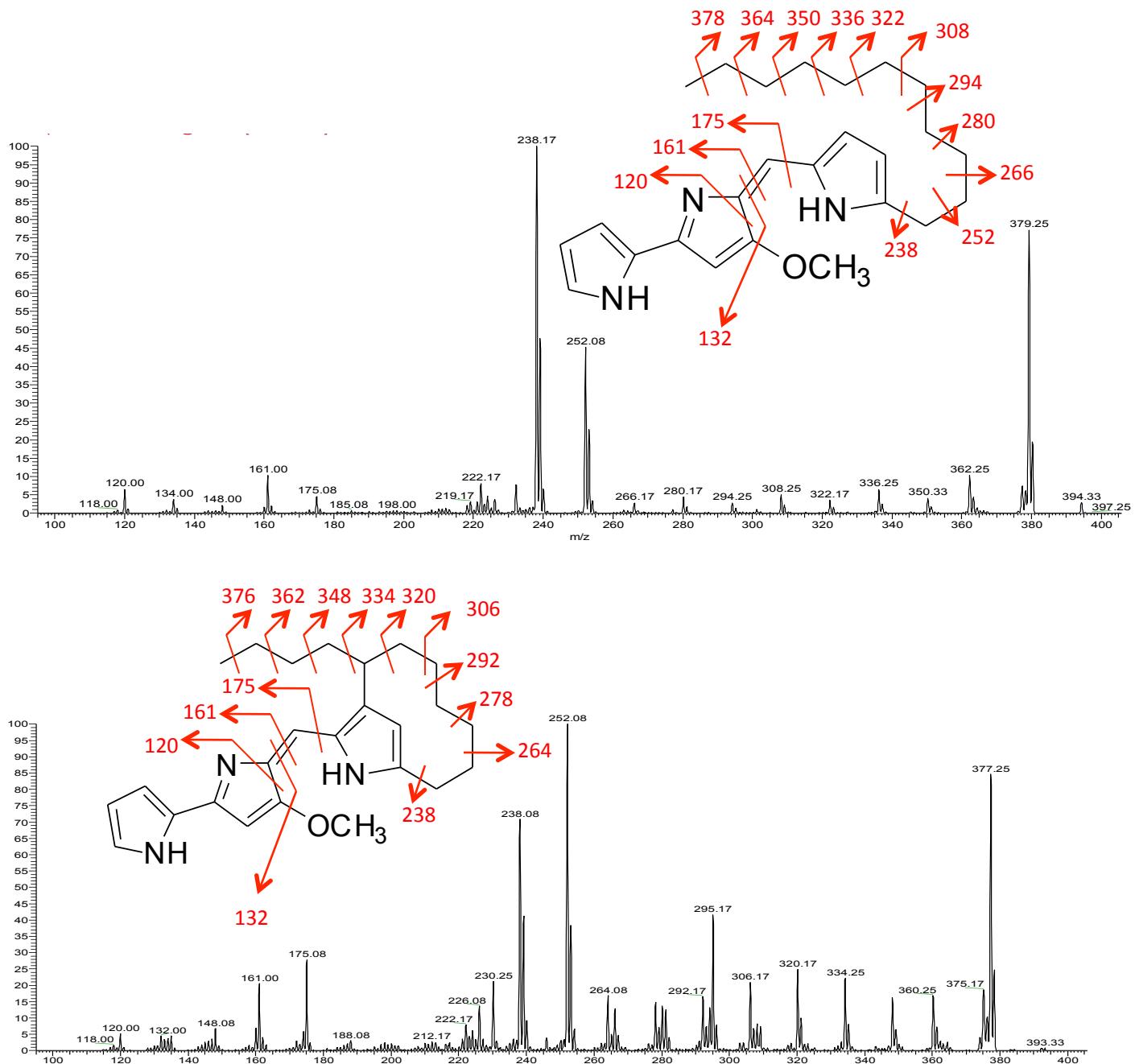
Streptomyces coelicolor A3(2): SapB



Description: CID MS2 for the $[M+2H]^{2+}$ ion of SapB. The spectrum is consistent with published MS2 data.⁶

Figure S4: Annotation of MS/MS spectra for reported compounds

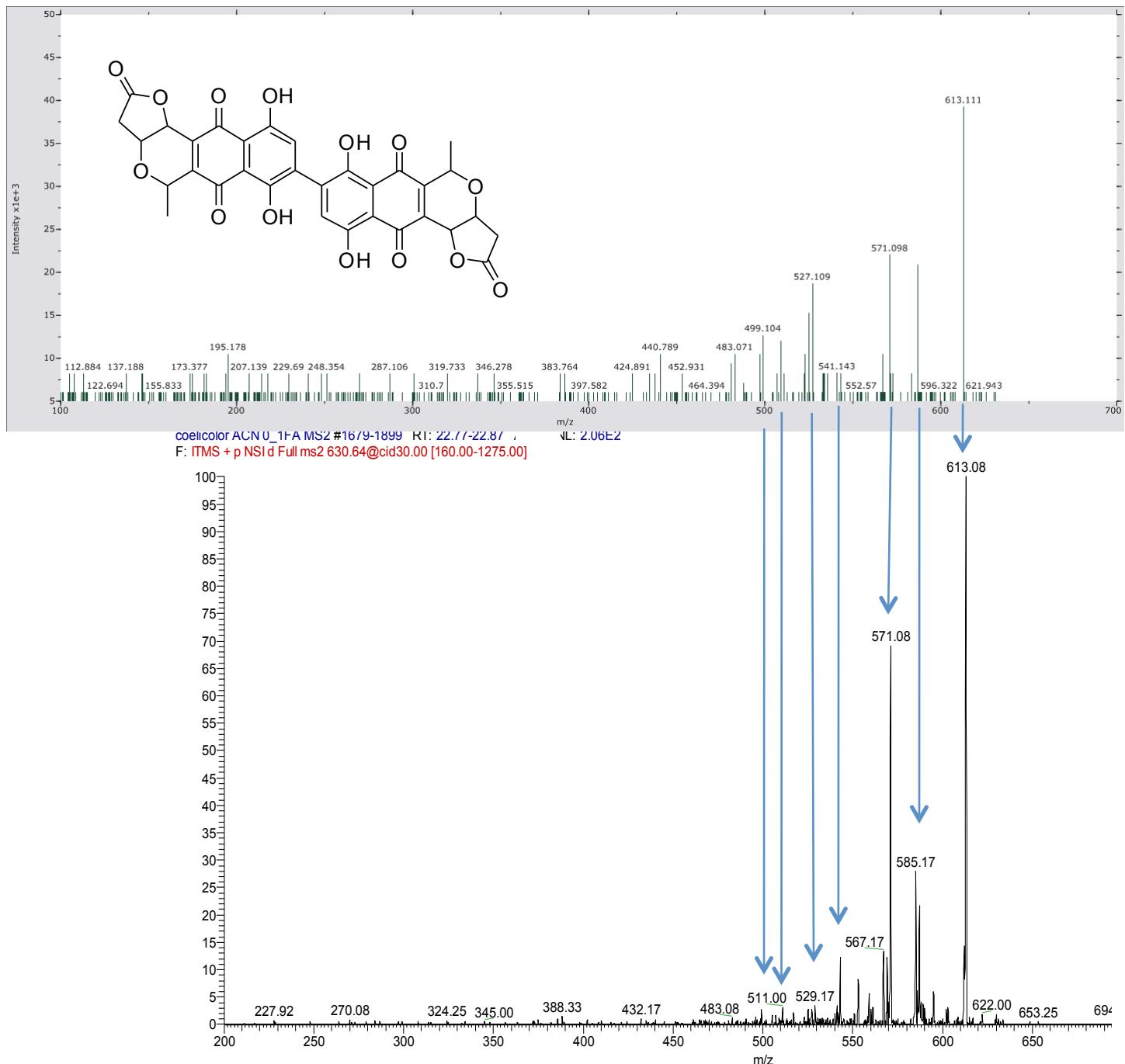
Streptomyces coelicolor A3(2): Prodiginine



Description: CID MS2 for the $[M+H]^+$ ions of cyclized and linear prodiginine (red pigments in *S. coelicolor*). The spectra are consistent with published MS2 data.⁷

Figure S4: Annotation of MS/MS spectra for reported compounds

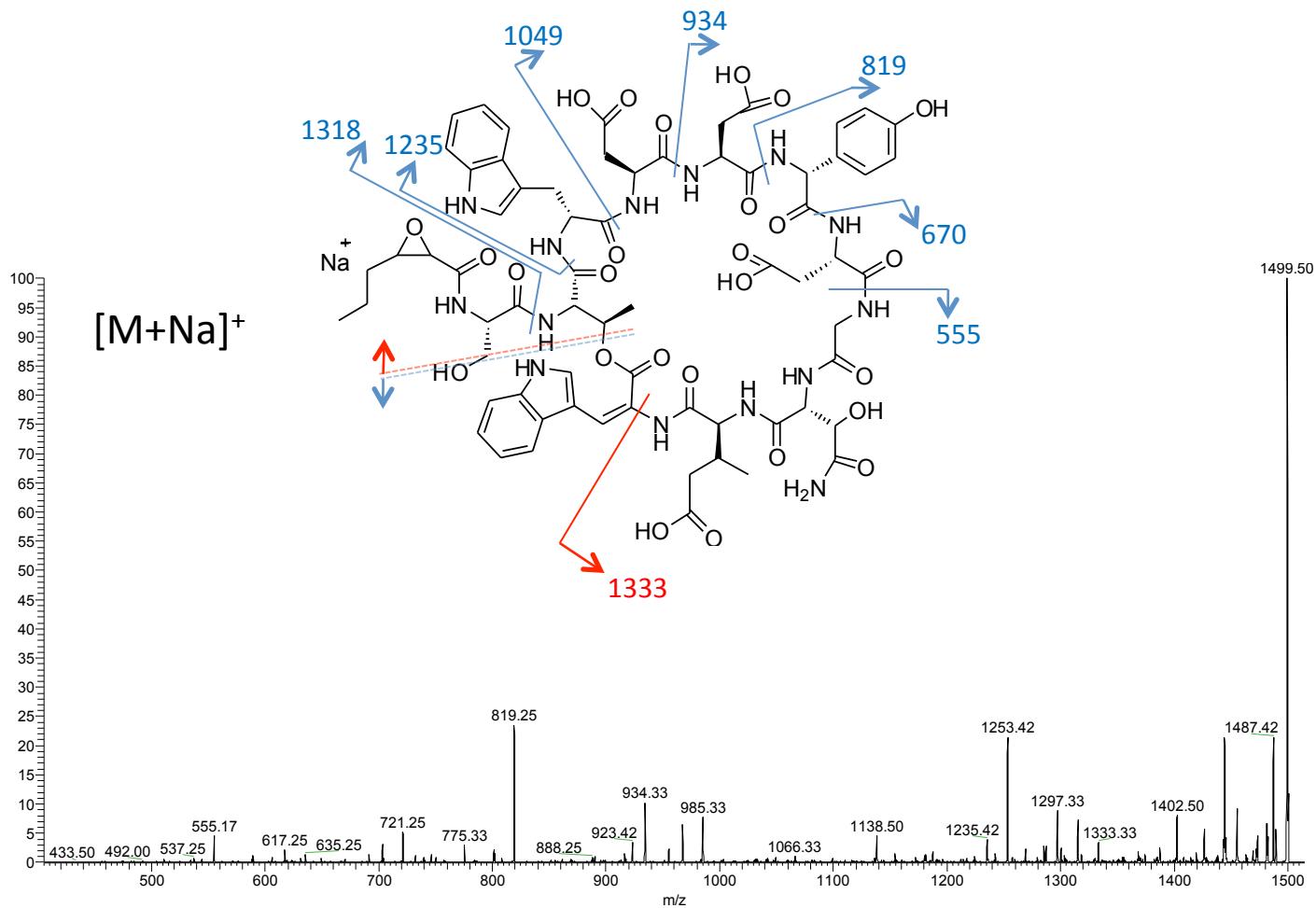
Streptomyces coelicolor A3(2): γ -Actinorhodin



Description: CID MS2 for the $[M+H]^+$ ion of γ -actinorhodin (the blue/purple pigment in *S. coelicolor*). The identity was confirmed by comparing the fragmentation pattern from the observed mass in the nanoDESI data set (bottom) with that of a standard of γ -actinorhodin (top). Blue arrows show identical peaks. We would like to thank Professor Richard van Breemen at University of Illinois at Chicago for donating the γ -actinorhodin reference spectrum.

Figure S4: Annotation of MS/MS spectra for reported compounds

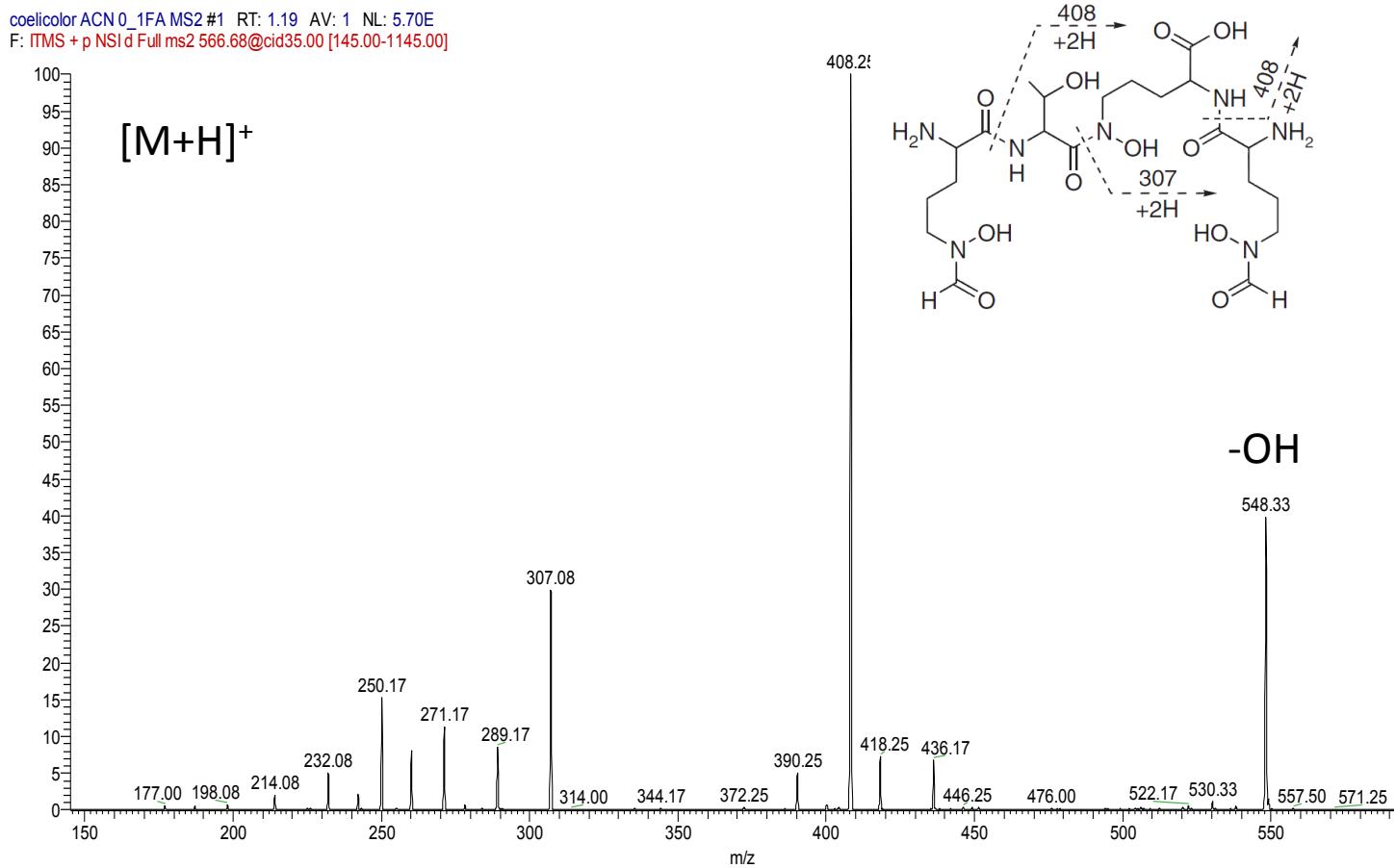
Streptomyces coelicolor A3(2): CDA



Description: Example MS2 CID spectra for the CDA family of compounds. The spectra are consistent with manual annotation as well as published MS2 spectra.⁸

Figure S4: Annotation of MS/MS spectra for reported compounds

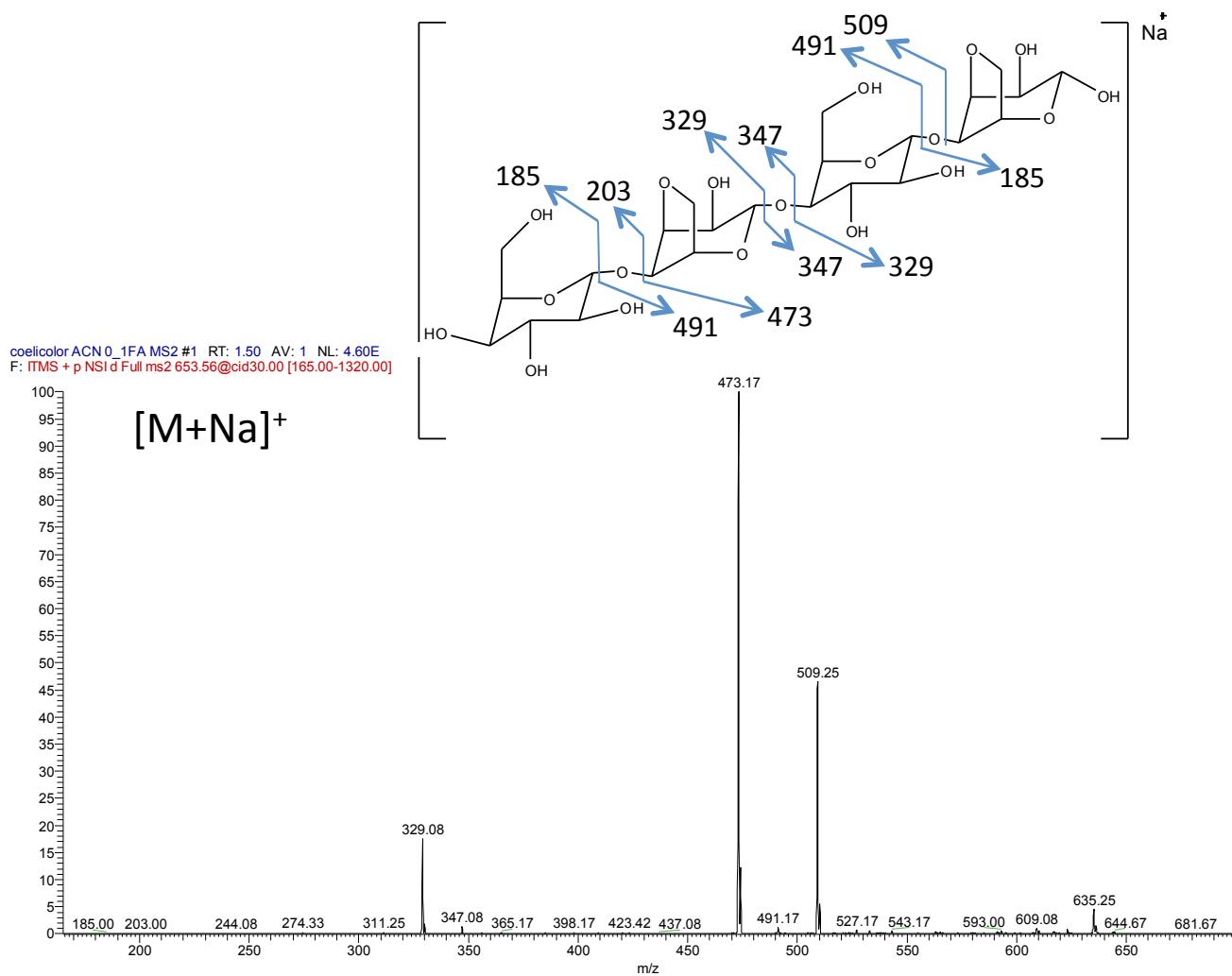
Streptomyces coelicolor A3(2): Coelichelin



Description: CID MS2 spectra for the [M+H]⁺ ion of coelichelin. The spectra is consistent with published MS2 spectra.⁹

Figure S4: Annotation of MS/MS spectra for reported compounds

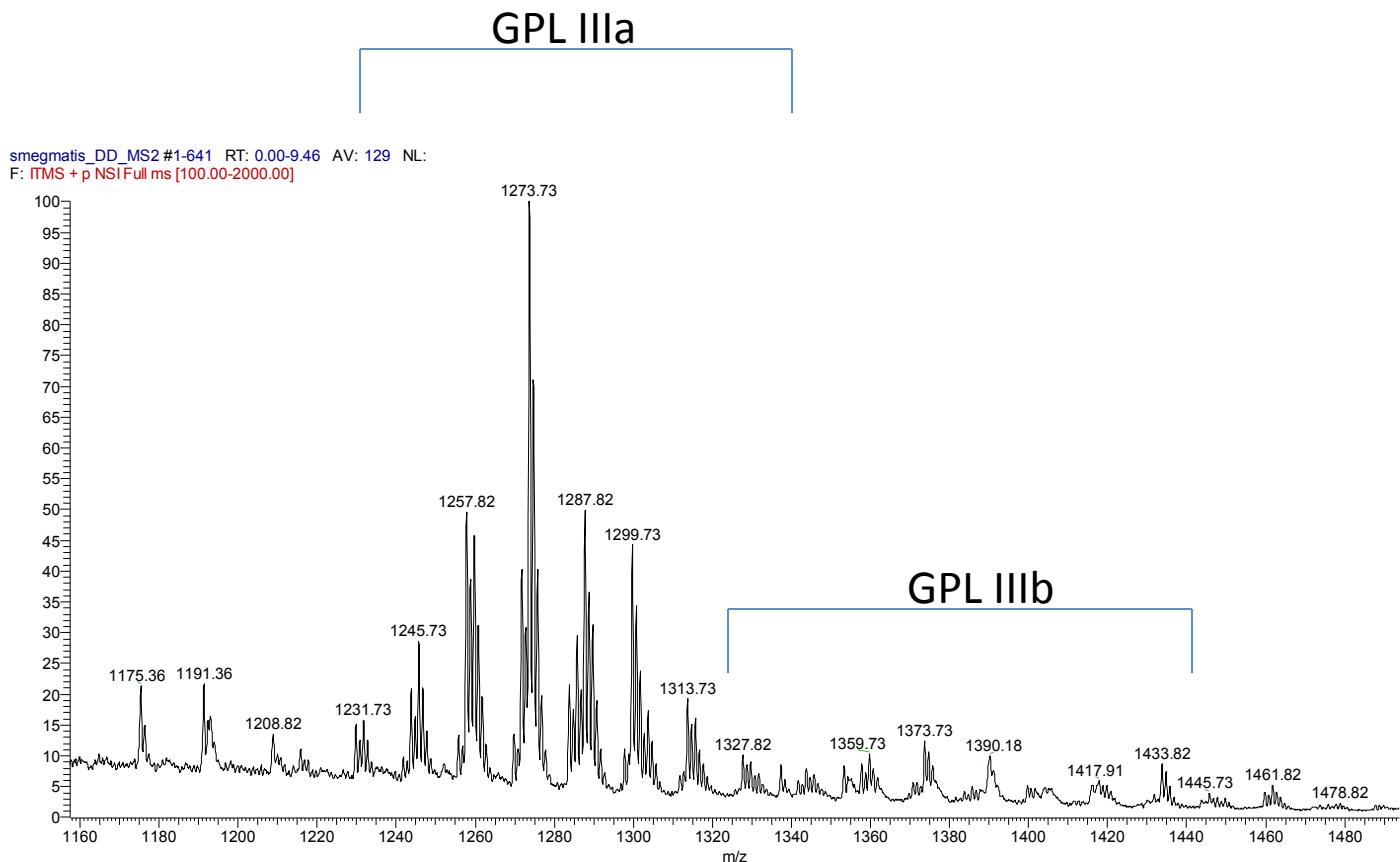
Example of agar polysaccharide



Description: Example subset of MS2 CID spectra for the polysaccharide found in the ISP2 group media. All MS2 spectra for the polymer show repeated losses of 162 and 180 Da due to shedding of hexose moieties in the chain. The spectrum is consistent with published MS2 data.¹⁰

Figure S4: Annotation of MS/MS spectra for reported compounds

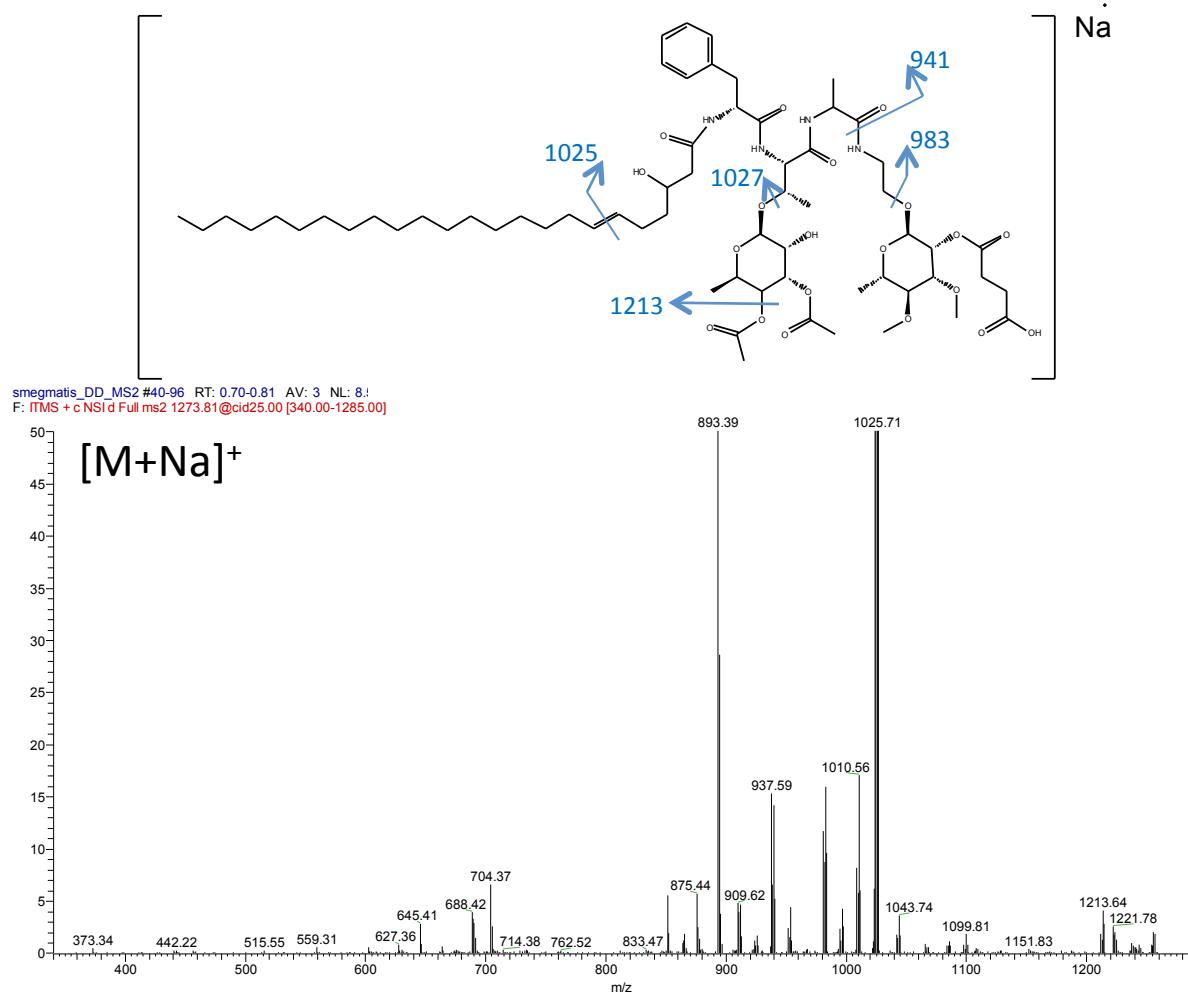
Mycobacterium smegmatis MC2: Glycopeptidolipids



Description: Annotation of the MS1 parent mass spectra for the glycopeptidolipids (GPL's) found in *M. smegmatis*. Peaks in the cluster are separated by 12 and 14Da which is indicative of a lipid chain. GPL III a and b are identical except for b containing one additional sugar unit.¹¹

Figure S4: Annotation of MS/MS spectra for reported compounds

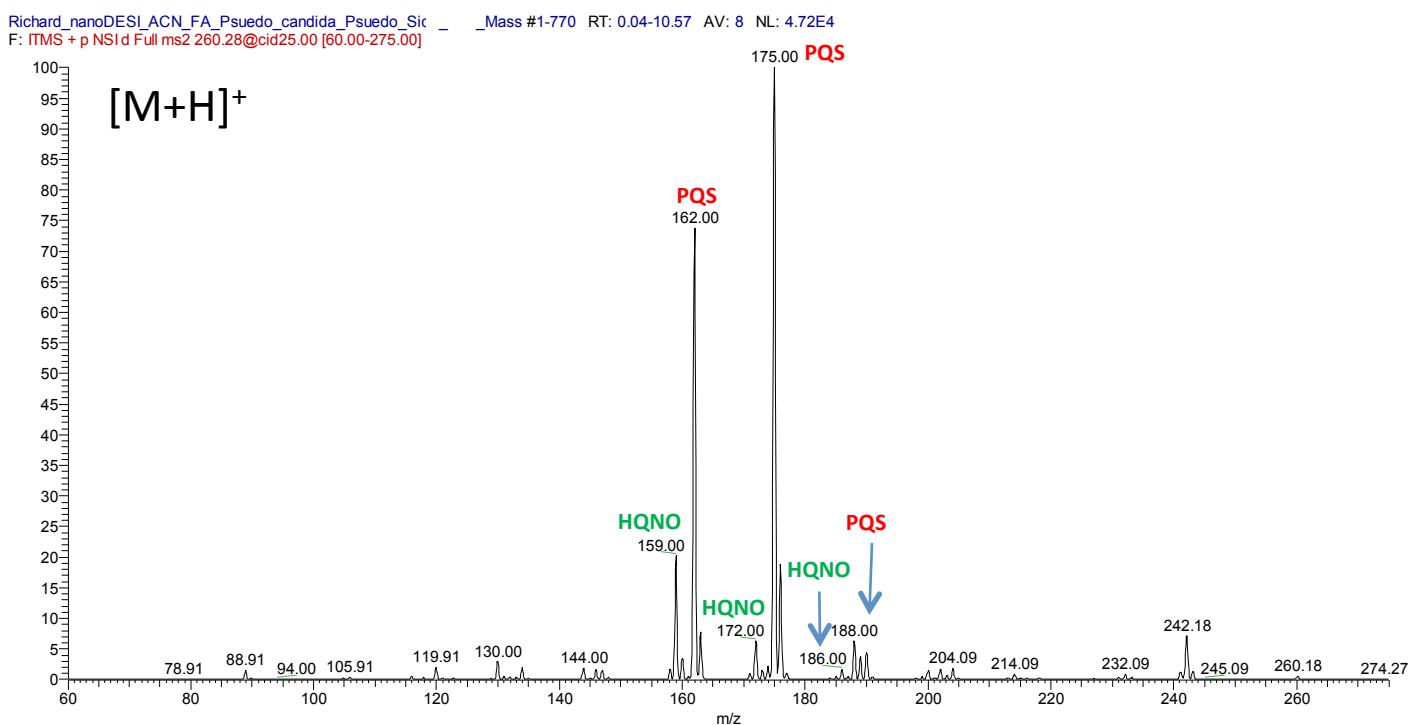
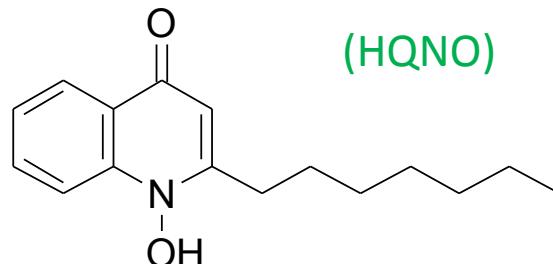
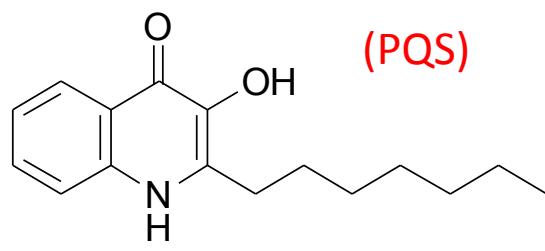
Mycobacterium smegmatis MC2: Glycopeptidolipids



Description: Example MS2 CID spectra for a GPL IIIa compound from the nanoDESI data. Data was confirmed by manual annotation as well as correlation so FAB MS2 published data.^{11,12}

Figure S4: Annotation of MS/MS spectra for reported compounds

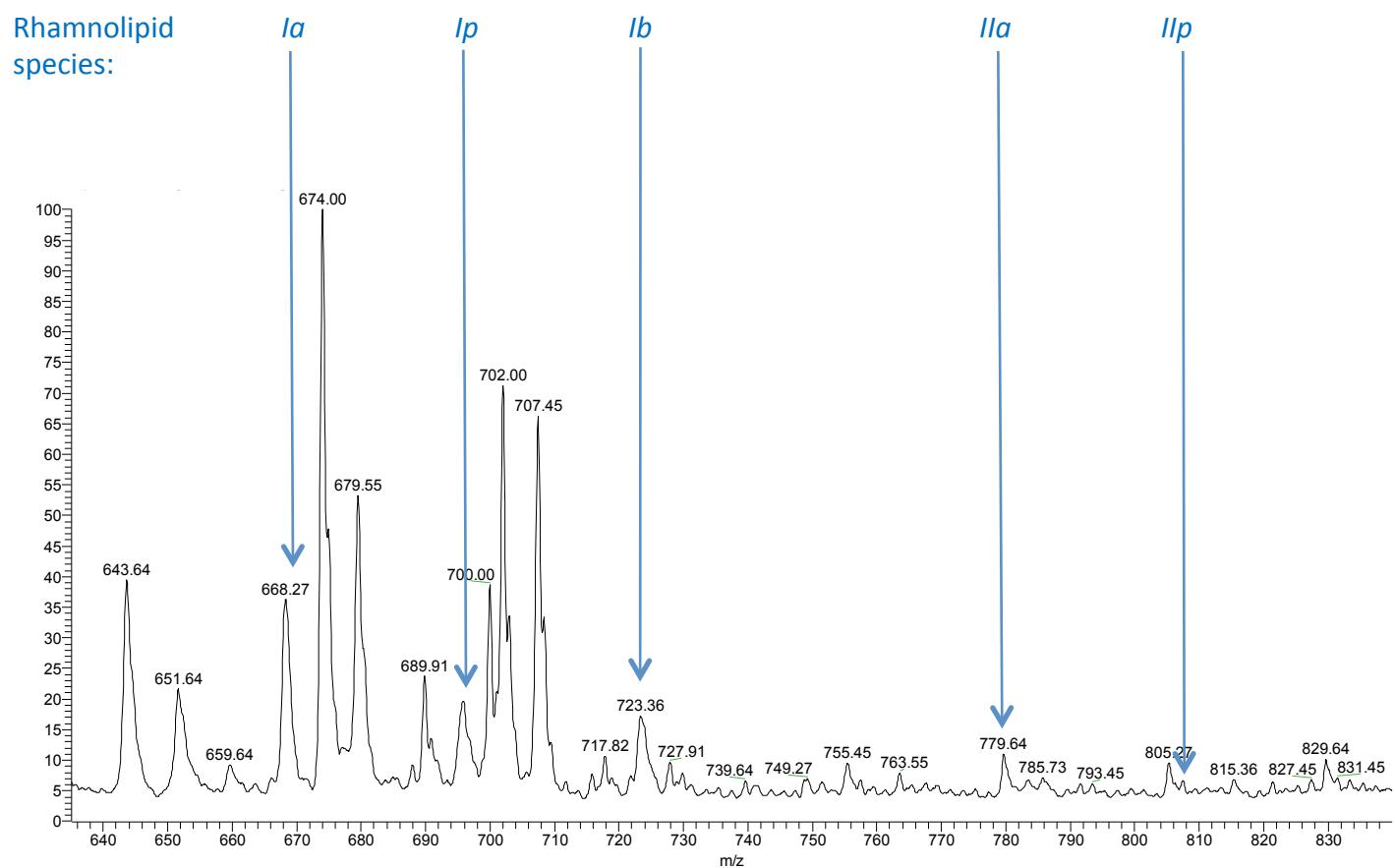
Pseudomonas aeruginosa PAO1: Quinolones



Description: Example MS2 CID spectra for the quinolone family of compounds from the nanoDESI data. The above spectrum shows a mixture of PQS and HQNO. The spectrum is consistent with published MS2 data for both compounds.¹³

Figure S4: Annotation of MS/MS spectra for reported compounds

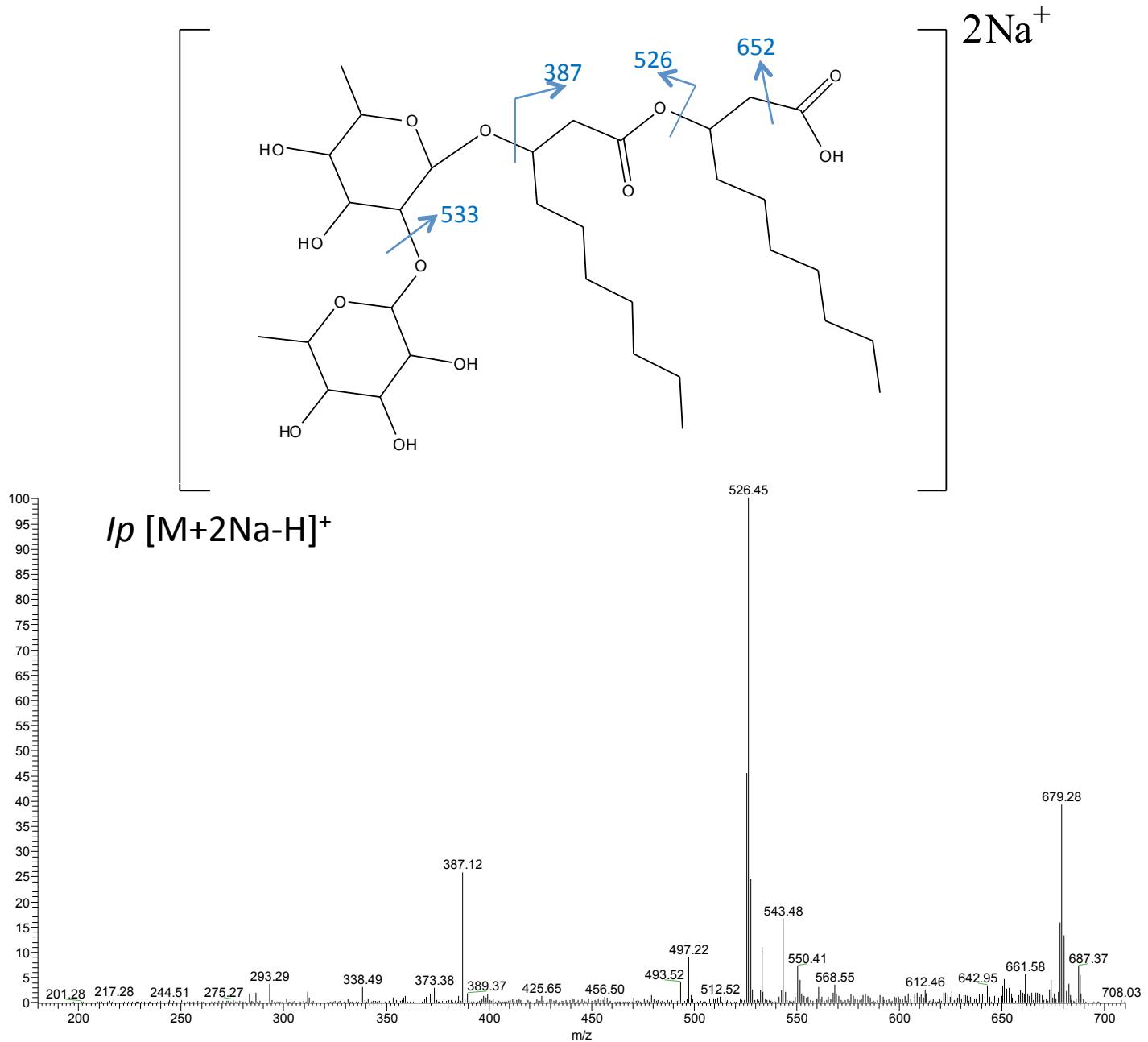
Pseudomonas aeruginosa PAO1: Rhamnolipids



Description: Annotation of the MS1 parent mass spectra for the rhamnolipid family of compounds. All ions are $[M+2Na-H]^+$.^{14,15}

Figure S4: Annotation of MS/MS spectra for reported compounds

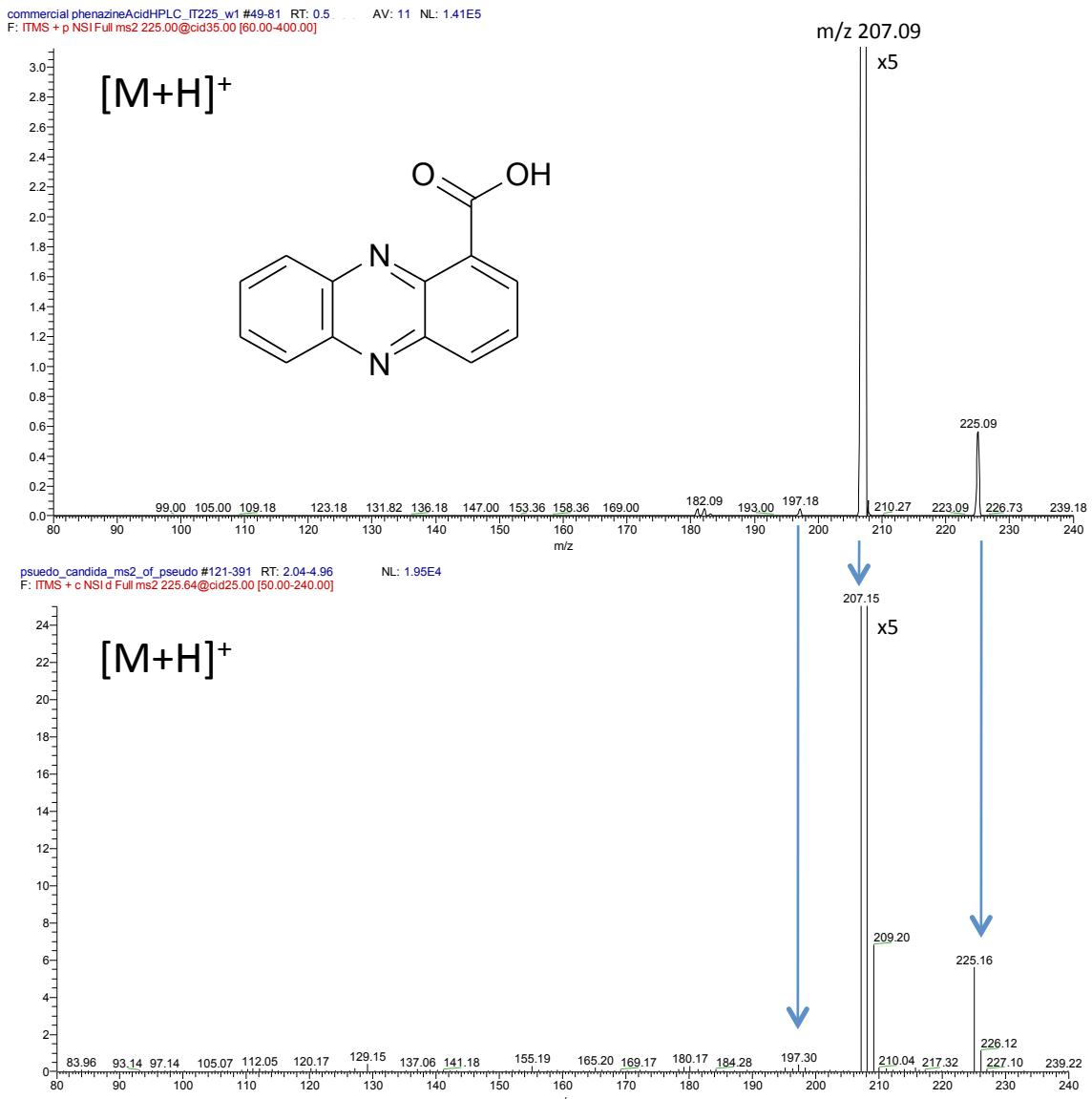
Pseudomonas aeruginosa PAO1: Rhamnolipids



Description: Example CID MS2 for the rhamnolipid family of compounds (Rlp). The MS2 fragmentation was consistent with published MS2 data and was confirmed by manual annotation.^{14,15}

Figure S4: Annotation of MS/MS spectra for reported compounds

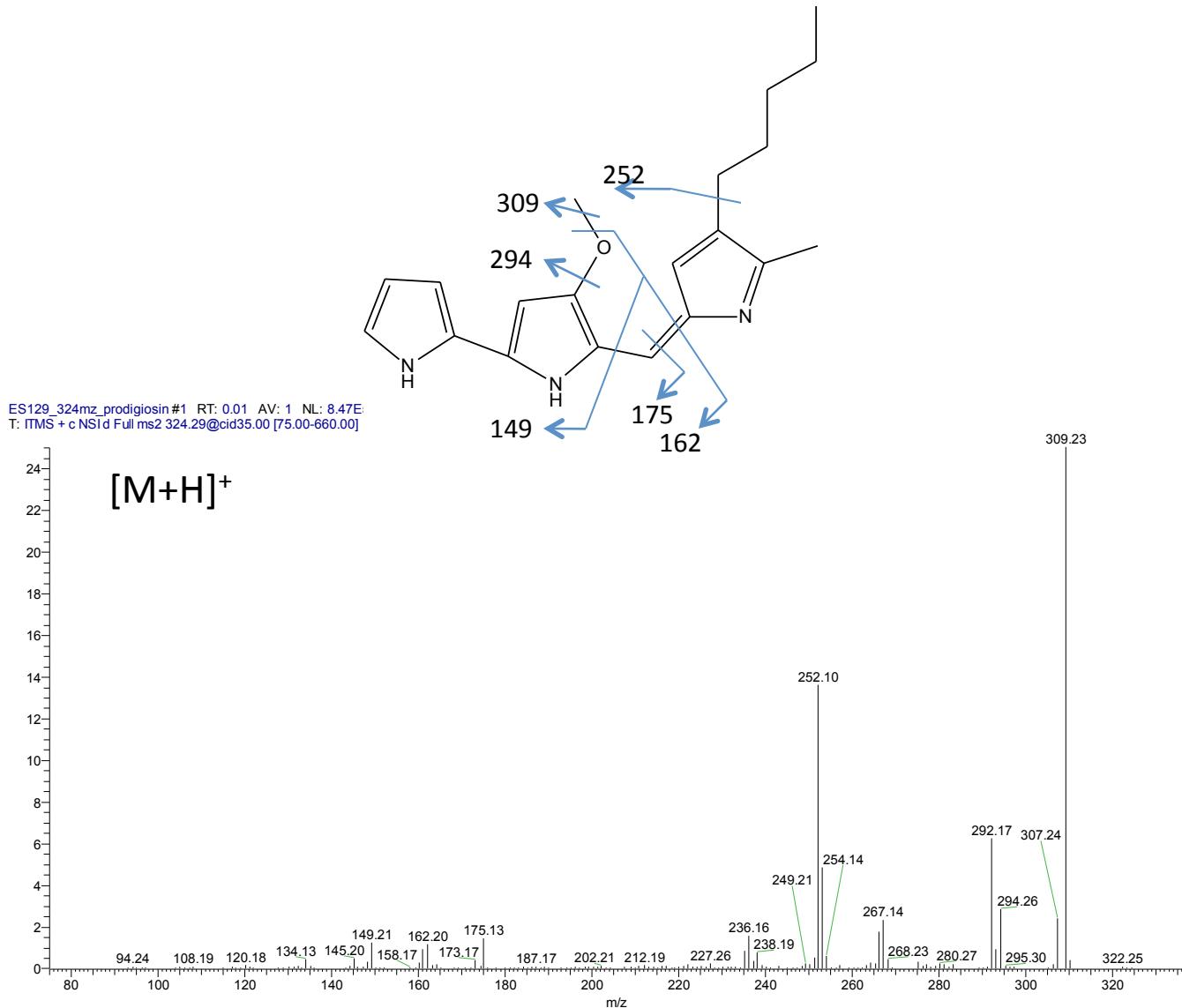
Pseudomonas aeruginosa PAO1: Phenazines



Description: Example CID MS2 for the phenazine family of compounds (phenazine-1-carboxylic acid shown). CID MS2 spectra from nanoDESI (bottom) is consistent with CID MS2 for commercial reference standard (top) as well as published data.¹⁶

Figure S4: Annotation of MS/MS spectra for reported compounds

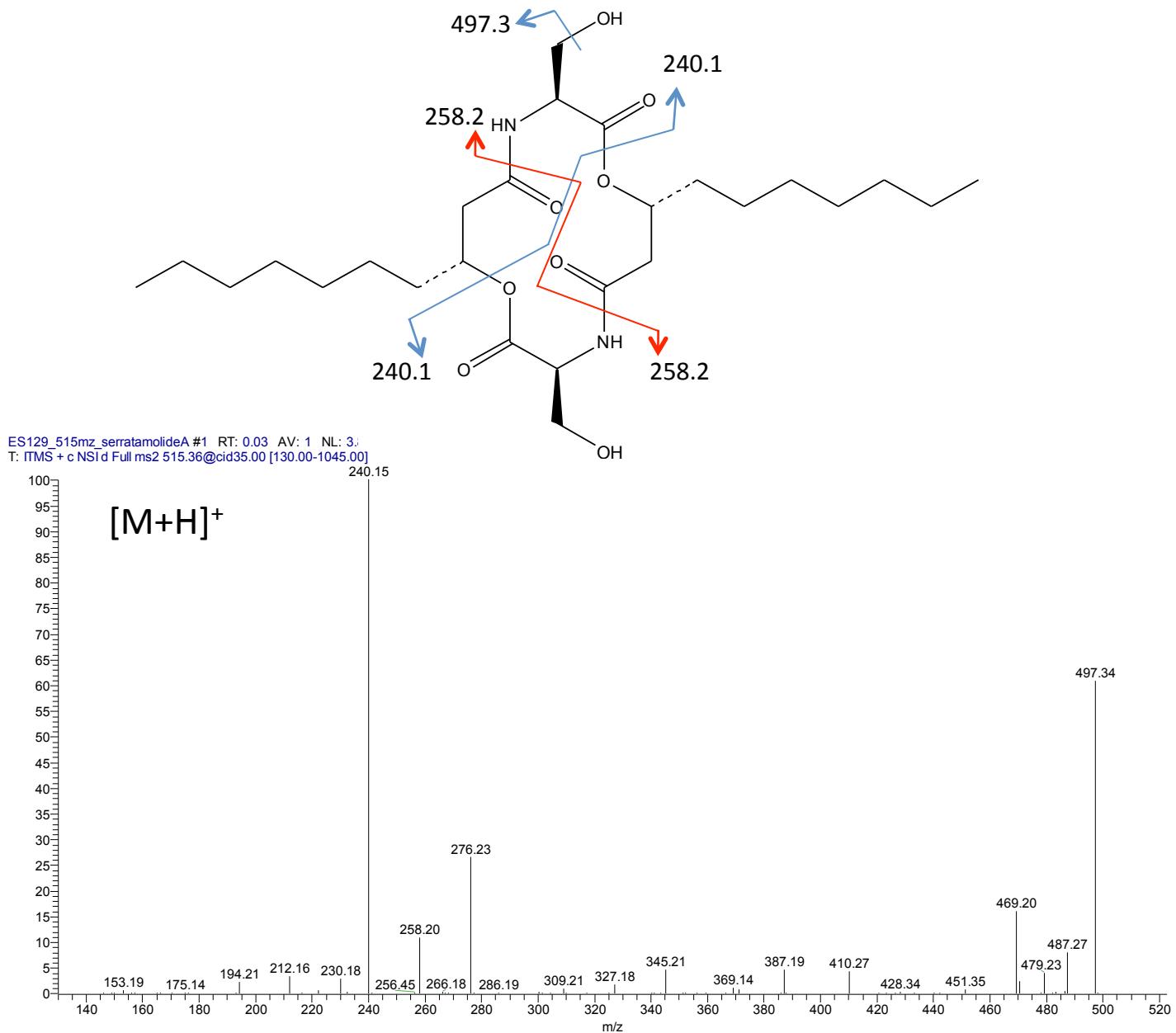
Serratia marcescens ES129: Prodigiosin



Description: Annotation of the CID MS2 for prodigiosin. Spectra is consistent with published MS2 data.
17

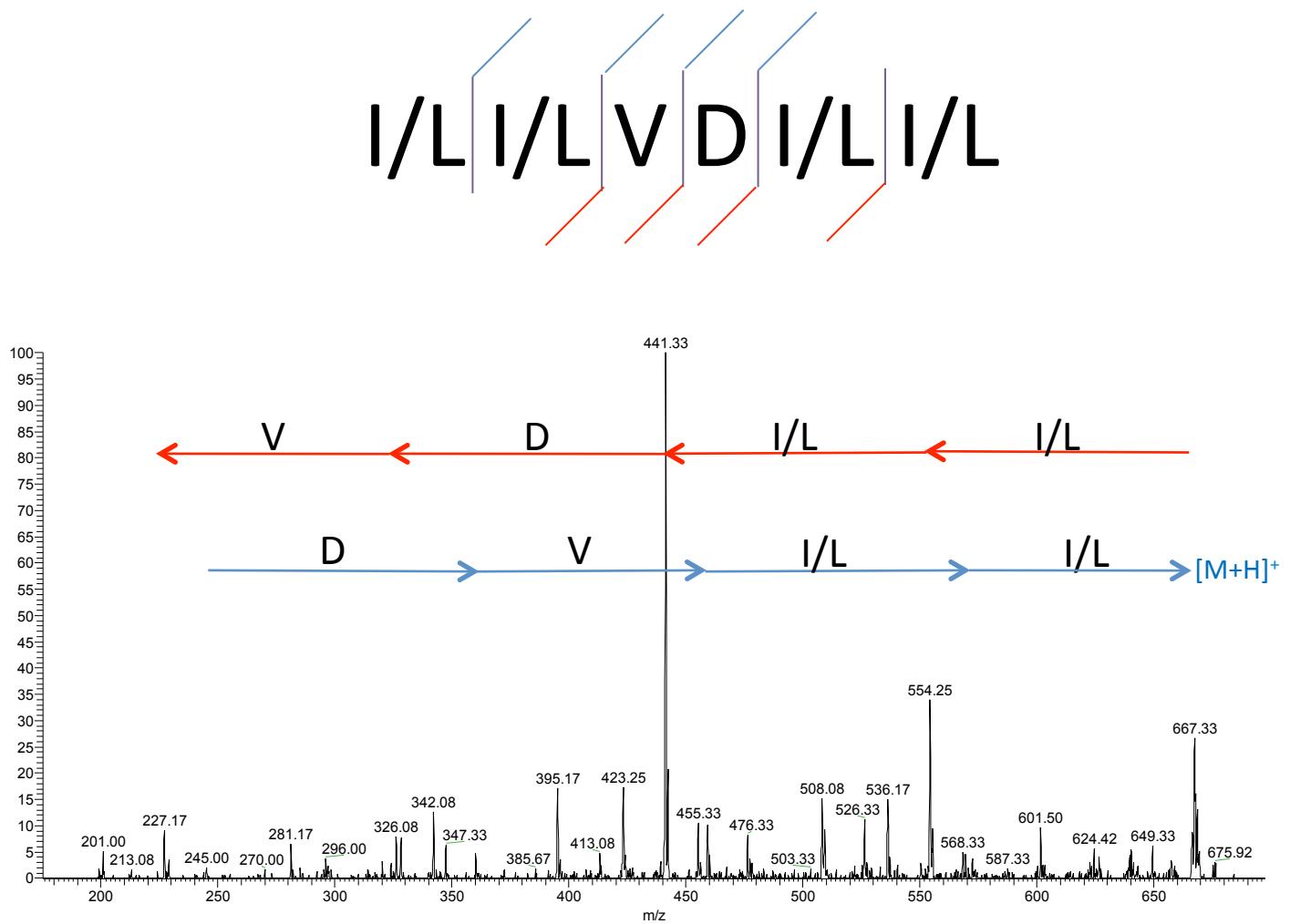
Figure S4: Annotation of MS/MS spectra for reported compounds

Serratia marcescens ES129: **Serrawettin W1/Serratamolide A**



Description: Annotation of the CID MS2 for serrawettin W1/Serratamolide A . Spectra is consistent with published MS2 data.¹⁸

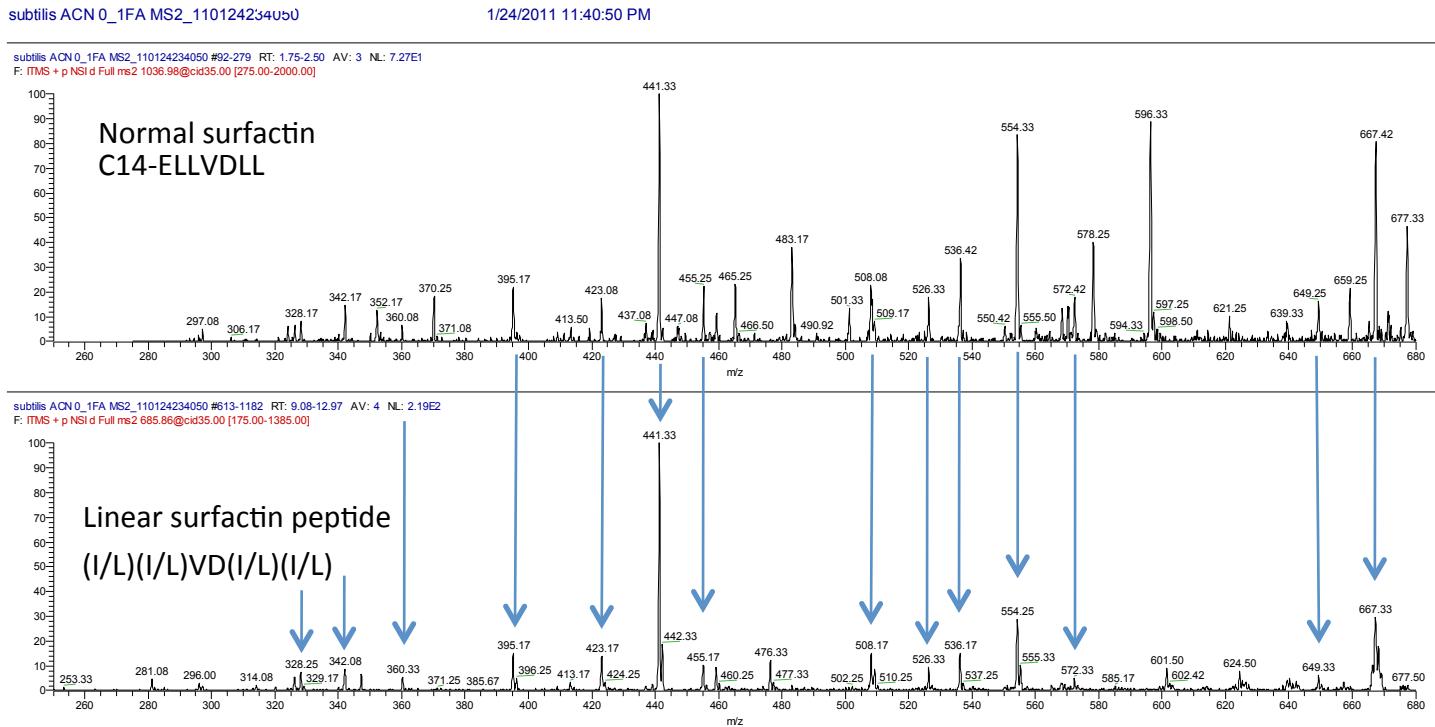
Figure S5: Annotation of MS/MS spectra for truncated linear surfactin



Calculated Mass: 685.4495 Da
 Observed Mass: 685.4512 Da

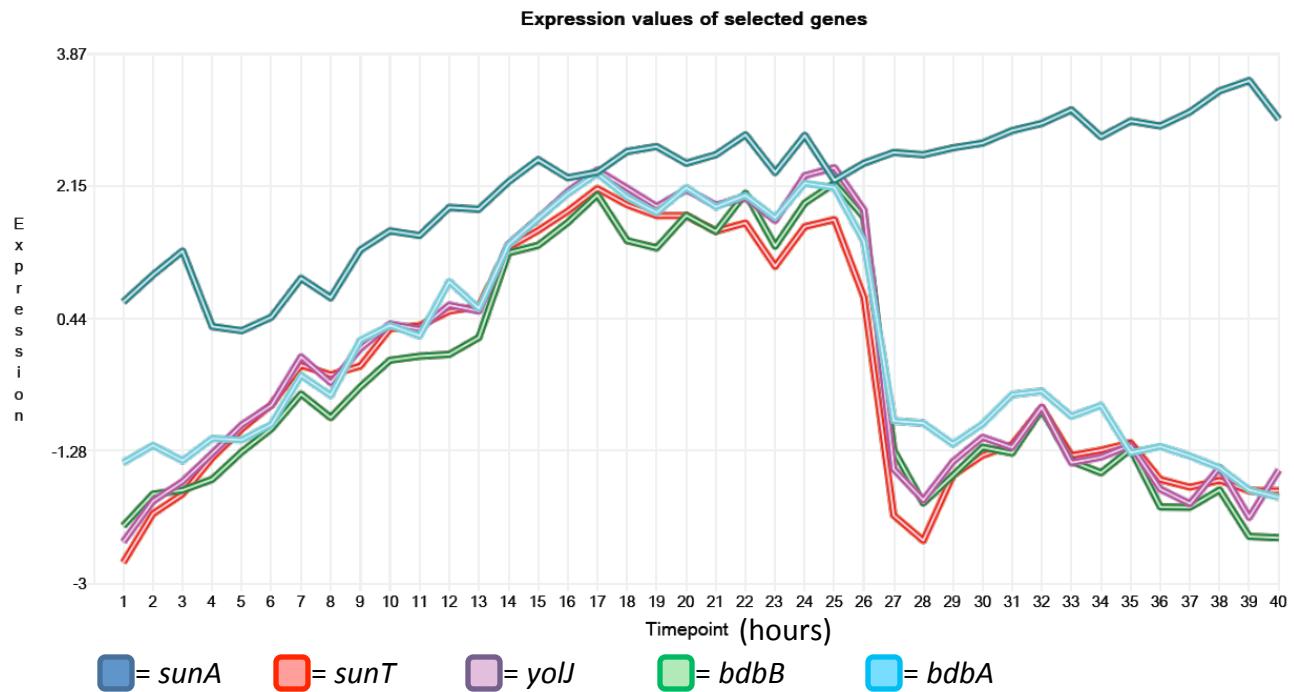
Description: Annotated MS2 CID spectra for the truncated linear surfactin peptide.

Figure S5: Annotation of MS/MS spectra for truncated linear surfactin

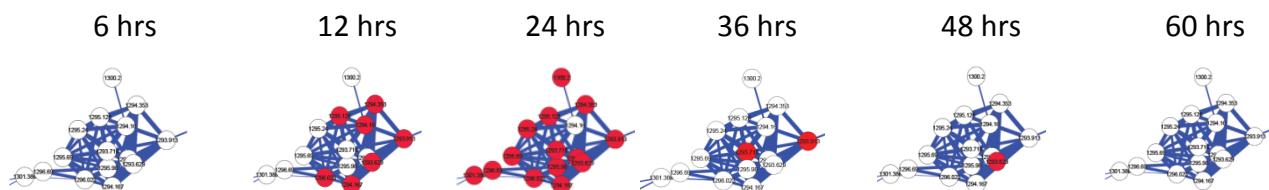


Description: Comparison of the CID fragmentation pattern of the linear surfactin peptide (bottom) with the that of the surfactin compound containing the same LLVDLL peptide sequence within the ring (top) shows identical daughter ion fragments. Note that only the lower mass fragments are displayed for the regular surfactin compound. Also, the reason for the peptide fragments for the cyclic peptide being equal to the fragments for the linear peptide is due to surfactin undergoing a ring opening upon fragmentation where the ester linkage is broken and the hydroxyl group is retained on the C-terminal side of the peptide.¹

Figure S6: RNA transcription time course for sublancin producing enzymes and correlation with MS/MS network profile

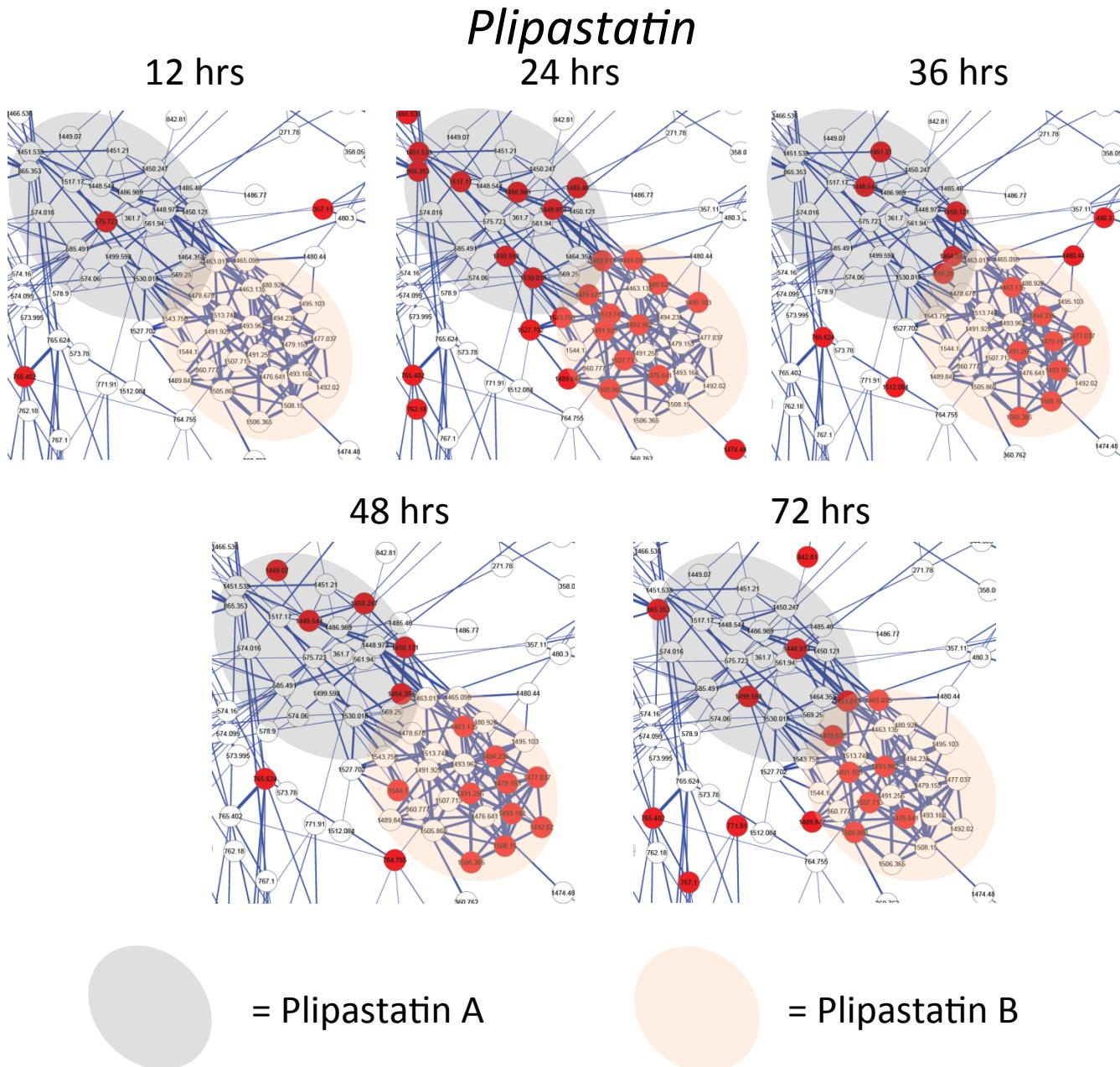


nanoDESI Time Course



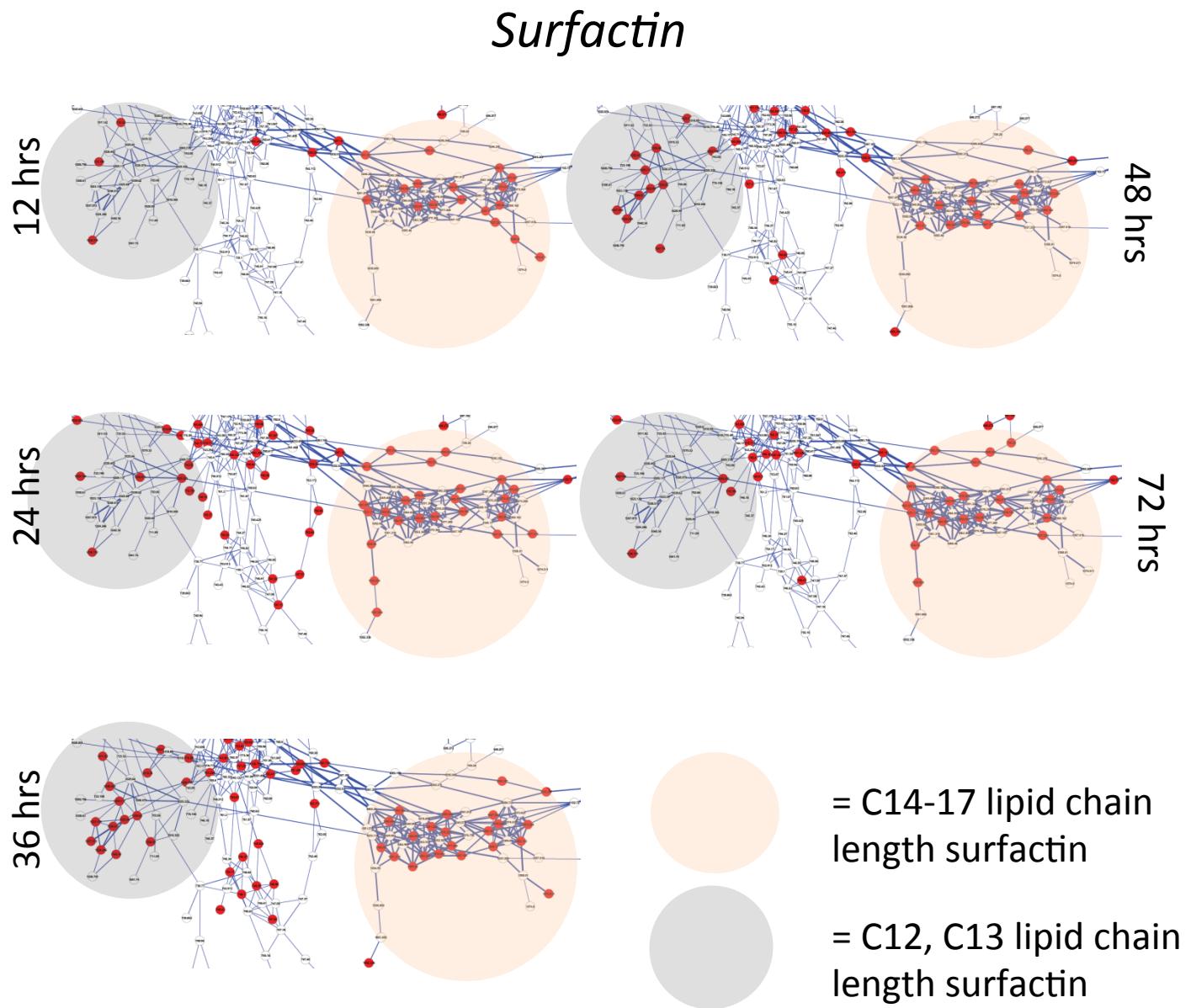
Description: Comparison of the RNA transcription data for sublancin biosynthetic machinery (top)¹⁹ with observed sublancin masses from the nanoDESI time course (bottom). The transcript data show a steep decline in transcription for the sublancin biosynthetic machinery after 25 hours (mainly the ABC transporter *sunT*, unknown gene *yolJ* and the two thiol-disulfide reductases *bdbA* and *bdbB* with the precursor gene *sunA* remaining relatively constant), which directly correlates to a steep decrease in observed sublancin masses within the nanoDESI time course from *B. subtilis* 3610. Red nodes in the time course indicate the mass was present at that time point. RNA transcript data plotted from <http://bacillus-subtilis168.appspot.com>.

Figure S7: *Bacillus subtilis* 3610 time course network annotation showing time dependent production of various species of surfactin and plipastatin



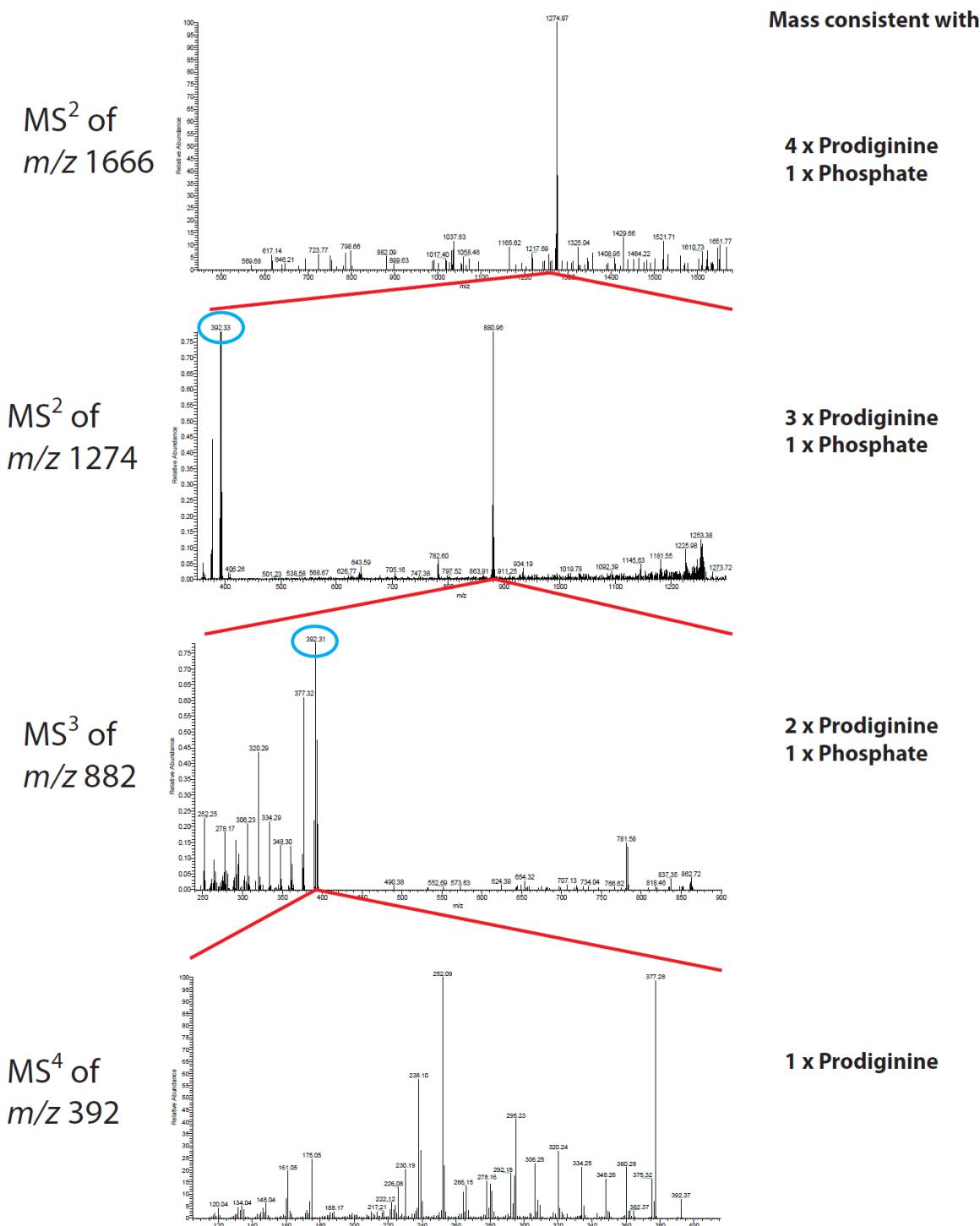
Description: When looking at the plipastatin node cluster within the molecular network for the nanoDESI time course of *B. subtilis* 3610, it was noticed that early on in the time course, plipastatin A (shorter lipid chain lengths) and plipastatin B (longer lipid chain lengths) were being made but at later time points only the longer lipid variants were observed in the mass spectra. This behavior also correlates to an overall increase in lipid production in the colony as observed by the networks after 36 hours. Red nodes indicate the mass was present.

Figure S8: *Bacillus subtilis* 3610 time course network annotation showing time dependent production of various species of surfactin and plipastatin



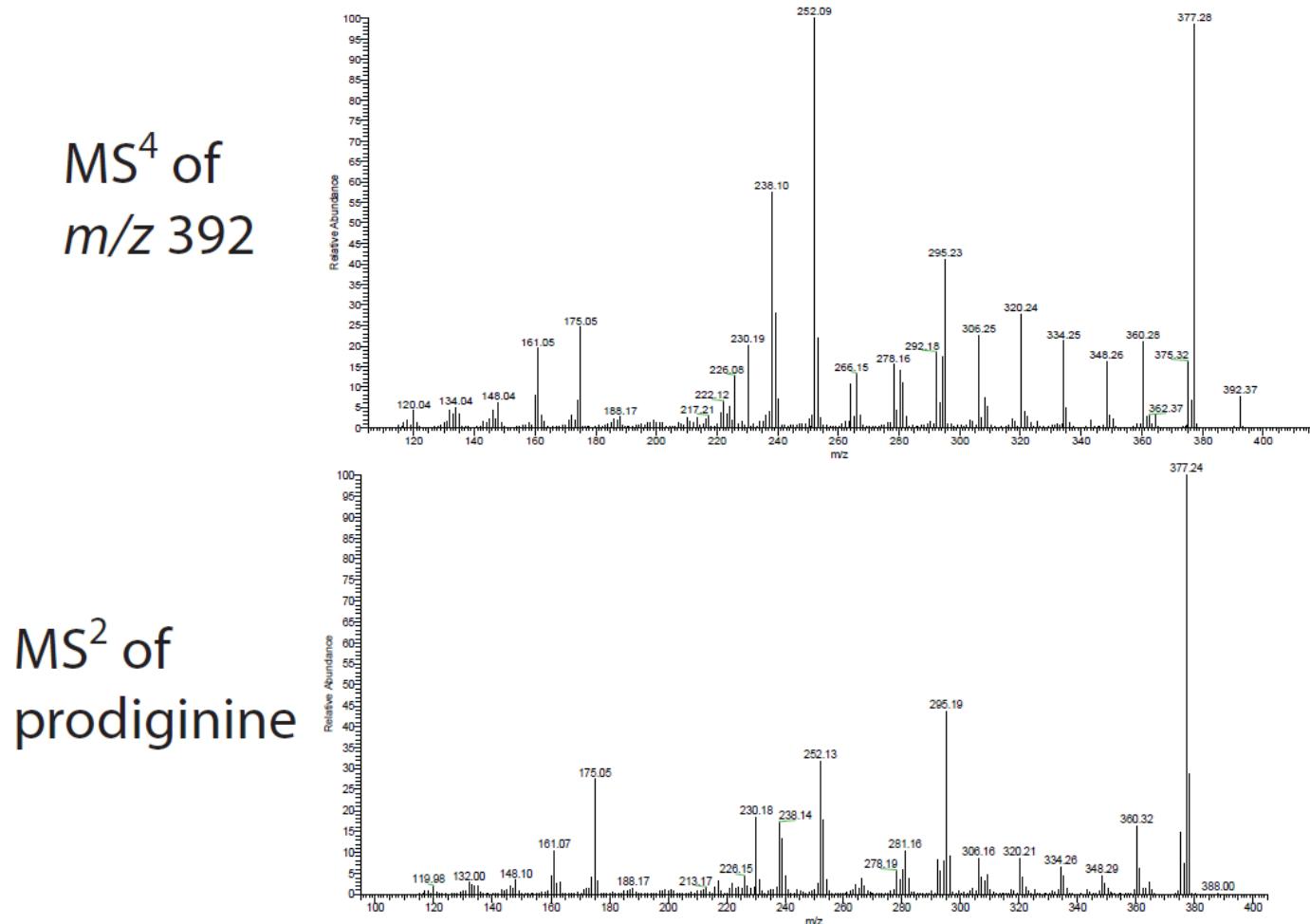
Description: When looking at the surfactin node cluster within the molecular network for the nanoDESI time course of *B. subtilis* 3610, it was noticed that while longer lipid lengths of surfactin A/B/C were being made throughout the time course, shorter lipid lengths (especially for surfactin C) were only produced within a time window of 24-48 hours after initial inoculation on the ISP2 agar plate

Figure S9: Mass spectra of prodiginine adducts/oligomers



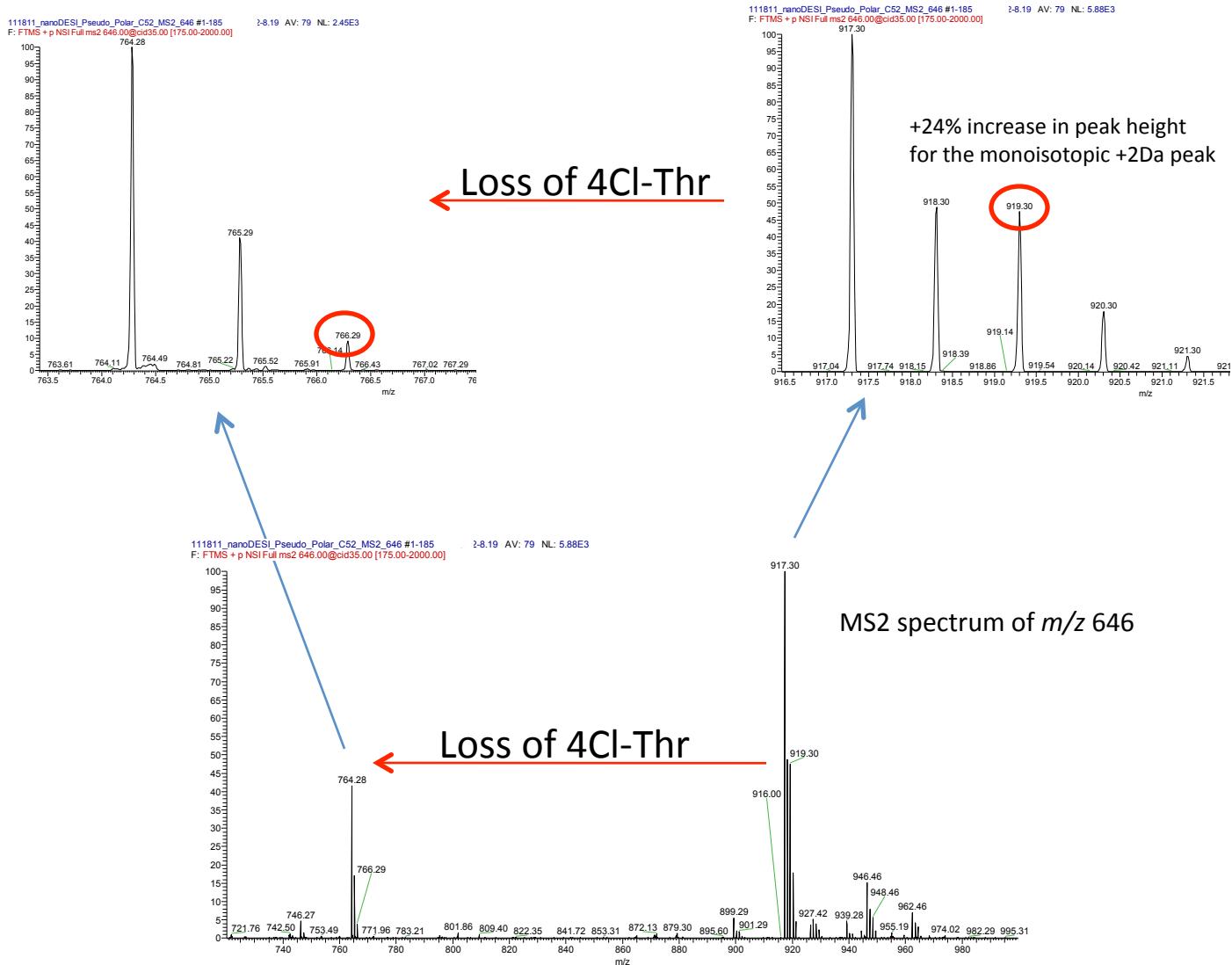
Description: Within the *S. coelicolor* colony that was interacting with the *B. subtilis* PY79 colony, an extremely high amount of cyclic prodiginine was observed in the spectra (much more than observed by MALDI and DESI previously in our lab). This was accompanied by many higher mass ions that contained prodiginine within their daughter ions indicating a possible polymer of the compound. MS2, MS3 and MS4 of one of these polymers revealed a possible oligomer of 2-4 prodiginine molecules with one phosphate.

Figure S9: Mass spectra of prodiginine adducts/oligomers



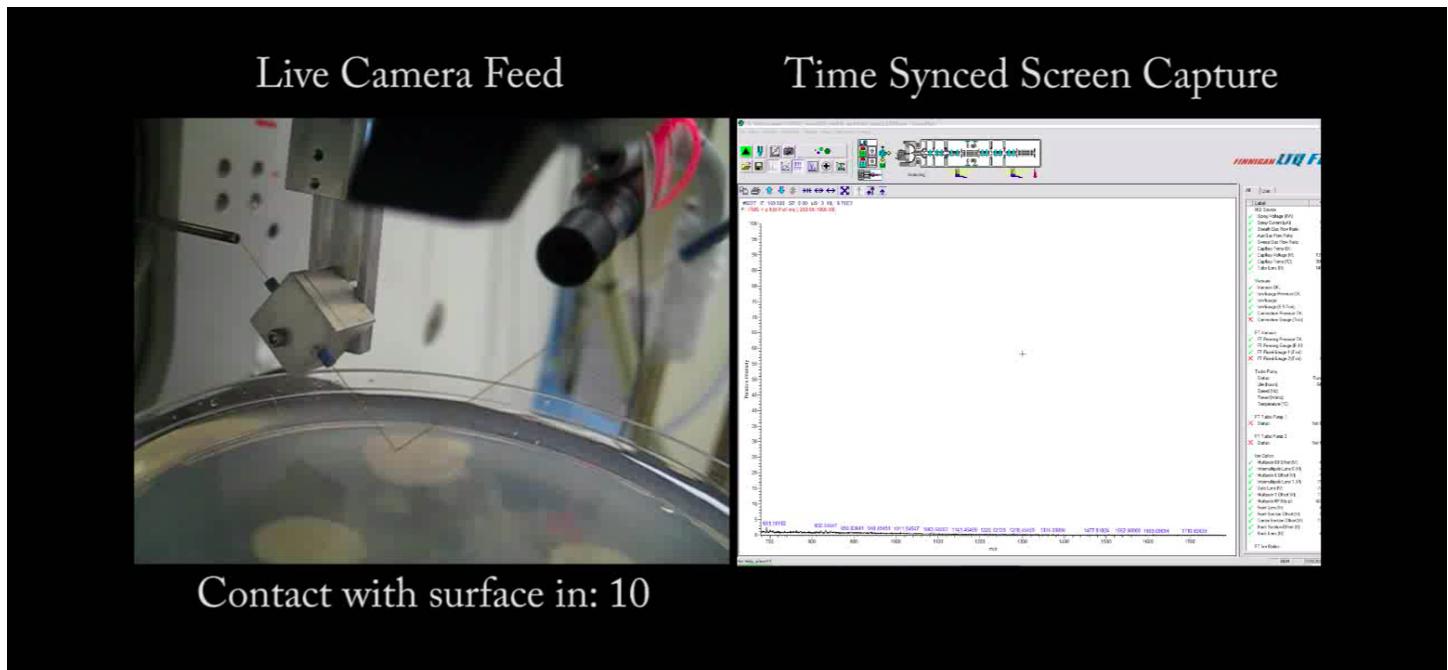
Description: To confirm that the mass of *m/z* 392 observed in the MS4 spectra of the oligomer was prodiginine (top), it was compared to the MS2 spectra of prodiginine to confirm its identity (bottom).

Figure S10: Example of chlorinated vs. non-chlorinated isotopic peak distributions



Description: Example of a typical chloronated (top right) and non-chloronated (top left) isotope distribution within the MS2 spectrum for thanamycin (MS2 of the +2 ion at m/z 646...bottom). Here we can clearly see that upon thanamycin losing the 4-chloro-threonine, the isotopic distribution of the mass signal changes from one indicative of halogenation (top right), where the monoisotopic peak +2 Da is 24% larger due to the presence of ^{37}Cl in addition to the typical 2x ^{13}C . The non-chloronated thanamycin fragment (top left) shows a typical carbon isotope distribution caused by the natural abundance of ^{13}C .

Video 1: This video shows the ionization and detection of microbial metabolites from living microbial colonies in real-time. This video can be found on the PNAS website.



References

1. Hue *et al.* (2001) Structural investigation of cyclic peptidolipids from *bacillus subtilis* by high-energy tandem mass spectrometry. *Rapid Comm. Mass Spectrom* 15, 203-209.
2. Gidden *et al.* (2009) Lipid compositions in *Escherichia coli* and *Bacillus subtilis* during growth as determine by MALDI-TOF and TOF/TOF mass spectrometry. *Int. J. Mass Spectrom.* 283, 178-184.
3. Pecci *et al.* (2010) LC/ESI-MS/MS characterization of lipopeptide surfactants produced by *Bacillus licheniformis* V9T14 strain. *J Mass Spectrom.* 45, 772-778.
4. Oman *et al.* (2011) Sublancin is not a lantibiotic but an O-linked glycopeptide. *Nat. Chem. Biol.* 7, 78-80.
5. Liu *et al.* (2010) Imaging mass spectrometry of intraspecies metabolic exchange revealed the cannibalistic factors of *Bacillus subtilis*. *PNAS* 107, 16286-16290.
6. Kodani *et al.* (2004) The SapB morphogen is a lantibiotic-like peptide derived from the product of the developmental gene *ramS* in *Streptomyces coelicolor*. *PNAS* 101, 11448-11453.
7. Gerber *et al.* (1975) Prodiginine (Prodiginin-like) pigments from *Streptoverticillium rubrifreticuli*, an organism that cause pink staining on polyvinyl chloride. *Appl. Microbiol.* 30, 807-810.
8. Hojati *et al.* (2002) Structure, biosynthetic origin and engineered biosynthesis of calcium dependent antibiotics from *Streptomyces coelicolor*. *Chemistry & Biology* 9, 1175-1187.
9. Lautru *et al.* (2005) Discovery of new peptide natural product by *Streptomyces coelicolor* genome mining. *Nat Chem Bio* 1, 265-269.
10. Wahl *et al.* (2010) Residual agar determination in bacterial spores by electrospray ionization mass spectrometry. *Anal. Chem.* 82, 1200-1206.
11. Villeneuve *et al.* (2003) Surface exposed glycopeptidolipids in *Mycobacterium smegmatis* specifically inhibit the phagocytosis of mycobacteria by human macrophages. *J. Biol. Chem.* 278, 51291-51300.
12. Mukherjee *et al.* (2005) Hyperglycosylation of glycopeptidolipid of *Mycobacterium smegmatis* under nutrient starvation: structural studies. *Microbiology* 151, 2385-2392.
13. Lepine *et al.* Electrospray/mass spectrometric identification and analysis of 4-hydroxy-2-alkylquinolines (HAQs) produced by *Pseudomonas Aeruginosa*. *J. Am. Soc. Mass Spectrom.* 15, 862-869 (2004).
14. Pajarron *et al.* (1993) Structure identification of natural rhamnolipid mixtures by fast atom bombardment tandem mass spectrometry. *Glycoconjugate J.* 10, 219-226.
15. de Koster *et al.* (1994) High-performance thin layer chromatography/fast atom bombardment (tandem) mass spectrometry of *Pseudomonas* rhamnolipids. *Biol. Mass Spectrom.* 23, 179-185.
16. Watson *et al.* (1988) Thermospray mass spectrometry analysis of phenazines. *Biomed. Environ. Mass Spectrom.* 17, 251-255.
17. Chen *et al.* (2008) Unusual odd electron fragments from eleven electron protonated prodiginine precursors using positive ion electrospray tandem mass spectrometry. *JASMS* 19, 1856-1866.
18. Dwivedi *et al.* (2008) Antimycobacterial serratamolides and diacyl peptoglucosamine derivatives from *Serratia* sp. *J. Nat. Prod.* 71, 637-641.
19. Ridder *et al.* (2011) Time-resolved transcriptomics and bioinformatic analyses reveal intrinsic stress responses during batch culture of *Bacillus subtilis*. *PLoS ONE*, 6.